

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2001, 17:58:39 ; Search time 1558.22 Seconds

(without alignments)
6352.317 Million cell updates/sec

Title: US-09-816-391A-1

Perfect score: 600
1 gctggcgcgcgccatga.....tattcgactagtcacgcg 600

Sequence: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Scoring table: 1472140 segs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_hgtgo_hum:*
31: em_hgtgo_inv:*
32: em_hgtgo_rnd:*
33: em_hgtgo_rnd:*
34: em_hgtgo_inv:*
35: em_hgtgo_rnd:*
36: em_hgtgo_rnd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	84	14.0	480	1	AB001381	AB001381 Streptomy
2	84	14.0	491	1	AF076983	AF076983 Caulobact
3	84	14.0	2327	1	AB033486	AB033486 Streptomy
4	84	14.0	35856	1	SCES9	AL118851 Streptomy
5	82.4	13.7	10960	1	AE005869	AE005869 Caulobact
6	75.2	12.5	20302	1	AF110185	AF110185 Burkholde
7	69.6	11.6	11691	1	AE004947	AE004947 Pseudomon
8	68.2	11.4	393	1	AF090448	AF090448 Rhizobium
9	67.4	11.2	13592	1	AE008054	AE008054 Agrobacte
10	66.8	11.1	970	1	AB032557	AB032557 Streptomy
11	66.8	11.1	42210	1	SCIC2	AL031154 Streptomy
12	66.4	11.1	11684	1	AE005903	AE005903 Caulobact
13	64.6	10.8	10851	1	AE004606	AE004606 Pseudomon
14	64.4	10.7	12116	1	AE004115	AE004115 Vibrio ch
15	63.8	10.6	3380	1	AF167159	AF167159 Sinorhizo
16	62.6	10.4	480	1	AF345628	AF345628 Pseudomon
17	61.8	10.3	568	1	ECHU2	X05994 E.coli HU2
18	61.8	10.3	11420	1	AE005631	AE005631 Escherich
19	61.8	10.3	18290	1	AE000473	AE000473 Escherich
20	61.8	10.3	176195	1	ECOTR89	U00006 E. coli chr
21	61.8	10.3	318703	1	AP002567	AP002567 Escherich
22	61.6	10.3	299350	1	SME591786	AL591786 Sinorhizo
23	61.2	10.2	489	1	AF146598	AF146598 Aeromonas
24	60.2	10.0	362	1	SM025149	U25149 Serratia ma
25	59.4	9.9	213732	1	AE001862	AE001862 Deinococc
26	59.2	9.9	2384	1	BBR132742	AJ132742 Bordetell
27	59.2	9.9	4439	1	BPFI32741	AJ132741 Bordetell
28	58	9.7	145911	1	AP003014	AP003014 Mesorhizo
29	56.8	9.5	309	1	BSHBSUOM	X6448 B. subtilis
30	56.8	9.5	312	12	SYNBACHSU	M38482 Synthetic B
31	56.4	9.4	485	1	AF345629	AF345629 Pseudomon
32	55.4	9.2	1374	1	MA297483	AJ297483 Myxococcu
33	55	9.2	359	1	SM025150	U25150 Serratia ma
34	53.8	9.0	584	1	BACHDBA	M73501 B. caldota
35	53.8	9.0	584	1	STRHUPA	M22975 S. typhimur
36	53.2	8.9	43339	1	STYSME1	AF170176 Salmonell
37	53.2	8.9	10135	1	AE006017	AE006017 Caulobact
38	52	8.7	478	1	STRHUP	X07844 Salmonella
39	50.6	8.4	270	1	BACHDBA	M73500 B. steatoth
40	50.6	8.4	315	1	AE0HUPA	M66574 Aeromonas P
41	50.6	8.4	429	1	BACDBPHU	D38080 Bacillus st
42	49.6	8.3	194140	1	AF242881	AF242881 Agrobacte
43	49.6	8.3	206479	1	AB016260	AB016260 Agrobacte
44	49.2	8.2	420	1	PSEHUPORA	L35257 Pseudomonas
45	49	8.2	270	1	BACHDB3	M73502 B. caldolyti

ALIGNMENTS

RESULT 1
AB001381
LOCUS AB001381 480 bp DNA BCT 16-JUN-1999
DEFINITION Streptomyces lividans DNA for HSI protein, complete cds.
ACCESSION AB001381
VERSION AB001381.1 GI:2104559
KEYWORDS HSI protein; hup; histone-like protein.
SOURCE Streptomyces lividans (strain:TK24) DNA.
ORGANISM Streptomyces lividans
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycetaceae; Streptomyces.
1 (sites)
Yokoyama, E., Doi, K. and Ogata, S.
Cloning and sequencing of the hup gene encoding the histone-like
protein HSI of Streptomyces lividans
JOURNAL Biochim. Biophys. Acta 1353 (2), 103-106 (1997)
MEDLINE 97438227
REFERENCE 2 (bases 1 to 480)
Yokoyama, E.
Direct Submission

(E-mail: y.koyama@agr.kyushu-u.ac.jp, Tel: 81-92-642-3059,
Fax: 81-92-642-3059)

FEATURES
Location/Qualifiers

1. .2327

/organism="Streptomyces lividans"

/strain="mk24"

/db_xref="taxon:1916"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

521. .802

/gene="hup"

helicase protein; integral-membrane protein; lysR-family
transcriptional regulator; malate oxidoreductase; membrane protein;
nitrate extrusion protein; oxidoreductase; RNA polymerase sigma
factor; secreted protein; serine/threonine protein kinase; small
protein B homologue; sugar hydrolase; sugar transporter inner
membrane protein; sugar transporter sugar-binding protein; tmRNA;
transcriptional regulator; transport protein; UDP-N-acetylglucosamine
transferase.

SOURCE
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

helicase protein; integral-membrane protein; lysR-family
transcriptional regulator; malate oxidoreductase; membrane protein;
nitrate extrusion protein; oxidoreductase; RNA polymerase sigma
factor; secreted protein; serine/threonine protein kinase; small
protein B homologue; sugar hydrolase; sugar transporter inner
membrane protein; sugar transporter sugar-binding protein; tmRNA;
transcriptional regulator; transport protein; UDP-N-acetylglucosamine
transferase.

SOURCE
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

helicase protein; integral-membrane protein; lysR-family
transcriptional regulator; malate oxidoreductase; membrane protein;
nitrate extrusion protein; oxidoreductase; RNA polymerase sigma
factor; secreted protein; serine/threonine protein kinase; small
protein B homologue; sugar hydrolase; sugar transporter inner
membrane protein; sugar transporter sugar-binding protein; tmRNA;
transcriptional regulator; transport protein; UDP-N-acetylglucosamine
transferase.

SOURCE
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

partial
/gene="SCE59_01c", possible oxidoreductase (partial), len:
/note="SCE59_01c, possible oxidoreductase (partial), len:
>116 aa; similar to N-terminal region of TR:O9Z8N9
(EMBL:AL034492) Streptomyces coelicolor putative
oxidoreductase subunit SC6C5_06, 296 aa; fasta scores:
opt: 223 z-score: 268.2 E(): 1.6e-07; 37.6% identity in 93
aa overlap and to TR:P95637 (EMBL:U65440) Rhodospirillum rubrum
palustris 4-hydroxybenzoyl-CoA reductase Hbd subunit, 327
aa; fasta scores: opt: 193 z-score: 233.2 E(): 1.4e-05;
42.6% identity in 94 aa overlap"
/codon_start=1
/transl_table=11
/product="putative oxidoreductase (partial)"
/protein_id="CAW72188.1"
/db_xref="GI:6911972"
/translation="MGOREGAVGLTTHAPQAAQVATLPTTIDEVAALTAIPAIVPA
GGTDMAANVSGLRPAALVGRISLRGMGYDQGMLLAGSLTHAMGRPDAALLI
PALAAARAGPPI"
complement(568..2205)
/gene="SCE59_02c"
complement(568..2205)
/gene="SCE59_02c"
/note="SCE59_02c, probable sugar hydrolase, len: 545 aa;
similar to TR:O9ZH39 (EMBL:AF072374) Pseudomonas sp.
S9 beta-N-acetylglucosaminidase Chlo, 783 aa; fasta
scores: opt: 1378 z-score: 1537.3 E(): 0; 41.6% identity
in 512 aa overlap and to SW:HEX1_VIBFU (EMBL:U041417)
Vibrio furnissii beta-hexosaminidase (EC 3.2.1.52) ExoI,
611 aa; fasta scores: opt: 1023 z-score: 1142.5 E(): 0;
40.0% identity in 487 aa overlap. Contains Pfam match to
entry PF00728 glyco_hydro_20, glycosyl hydrolase family
20"
/codon_start=1
/transl_table=11
/product="putative sugar hydrolase"
/protein_id="CAW72189.1"
/db_xref="GI:6911973"
/translation="MTDLTVYDANDVIPPRHAAVADACFDLPRTLVADDEAGHT
EKMRLATGAATGLPAPEGGGGVARSLSIDELADEXYRLDIADVGRLGGPAGI
FMAQTOLRIQLGPAPRRAPICRGWRRLPLIGENAPRGKMLDVSRMPKDGVI
LNQDLMHAHKNTLVHFHLITDDQGRLEIKRPKITGTGSRSRKTRPSRWEDERK
PHGGYTQCDIREIVAATAARIITVLPELDVPSSQAIAVPELGNDVIDTALS
VNWTCGSPVDLAFTENTLRETFGEVEVELPEVFEVIGDECEKQDMWSAQAQR
MAELGIDDELQSQVFIOHFDTWLAGRBRLLIGMEILLEGLAGKAASVRGAGGII
AABAGHDVVMDPEOVVLTNFERDGGDEBPVIAVPTLEDVYREPVPATLPESGS
HYTGCAANTTEVTNOERYDTLPRLAAPFAVMASLPPADRDVAAFERRMATHY
RLIDLAAVARPPAGRPFRPQRRPGVLRPMQDPFPNK"
complement(1096..1803)
/gene="SCE59_02c"
complement(1096..1803)
/note="Pfam match to entry PF00728 glyco_hydro_20,
glycosyl hydrolase family 20, score 199.70, E-value
4.4e-56"
complement(2210..3055)
/gene="SCE59_03c"
complement(2210..3055)
/gene="SCE59_03c"
/note="SCE59_03c, possible sugar transporter inner
membrane protein, len: 281 aa; similar to TR:CAB59690
(EMBL:A1J3262) Streptomyces coelicolor putative sugar
transporter inner membrane protein SCP11_13, 285 aa; fasta
scores: opt: 730 z-score: 883.7 E(): 0; 42.1% identity in
278 aa overlap and to TR:NAFL1008 (EMBL:AE001988)
Deinococcus radiodurans ABC transporter, permease protein,
MalP family DRI436, 283 aa; fasta scores: opt: 645
z-score: 781.7 E(): 0; 38.2% identity in 272 aa overlap.
Contains Pfam match to entry PF00528 BPD_transp.
Binding-protein-dependent transport systems inner membrane
component and match to Prosite entry PS00402
Binding-protein-dependent transport systems inner membrane
comp sign. Contains possible hydrophobic membrane spanning
regions"
/codon_start=1

		/transl_table=1	
		/product="putative inner sugar transporter inner membrane protein"	
		/protein_id="CA872190.1"	
		/db_xref="GI:6911974"	
		/translation="MNLRLARVRPRMLAEVSLALLAVVAFLYMWLVSAFPAAGEE IESEPPWTLTPTSDISDFRVEEQEGREINSLIYAGSVIVASIALAFAAVRIR FRFRRTLLIMFLVQMVPVALTIPLFMRFQGLNLSILHIFAIPELAIM MGEFYKAVPEALPEEAATIDGASTRRPKNQITLPVLVPGLYATVSFEFSIRNDPLRA KSFISTOSTSOTLEPMALVFYKPDPPDGMVMASTWTIPVLVEFVLQRRLVSQSGA VAND"	
misc.feature		complement(2342..2563)	
		/gene="SCB59_03c"	
		/note="pfam match to entry PF00528 BPD transp. Binding-protein-dependent transport systems inner membrane component, score 58.90, E-value 1.1e-13"	
misc.feature		complement(2474..2560)	
		/gene="SCB59_03c"	
		/note="PS00402 Binding-protein-dependent transport systems inner membrane comp sign"	
		complement(3052..4002)	
		/gene="SCB59_04c"	
CDS		complement(3052..4002)	
		/gene="SCB59_04c"	
gene		/note="SCB59_04c, possble inner sugar transporter inner membrane protein, len: 316 aa; similar to TR:CB59589 (EMBL:AL132662) Streptomyces coelicolor putative sugar transporter inner membrane protein SCF1.12, 302 aa; fasta scores: and to TR:CB59533 (EMBL:M117322) Streptomyces coelicolor putative binding-protein-dependent transport protein SCF1.13, 317 aa; fasta scores: opt: 527 z-score: 611.6 (E): 1.2e-26; 33.1% identity in 302 aa overlap. Contains pfam match to entry PF00528 BPD transp. Binding-protein-dependent transport systems inner membrane component. Contains also possible hydrophobic membrane spanning regions"	
		/codon_start=-1	
		/transl_table=-1	
		/product="putative inner sugar transporter inner membrane protein"	
		/protein_id="CA872191.1"	
		/db_xref="GI:6911975"	
		/translation="MSSGGVRCGRVRRRRLSRSSPPRGSSSTPWLYIALPALVLG GLVTPYVOGLISFILETOAOVSGSPAPFOGGAANEELFGDEFNRVLATVFAA ACVATLAVGCALAVALITRVRAVRLMALMGALCANAPATVGSTMILLDEPDFGV NRMLGDHSMTGYRSYAFLLILEVVMSPPFMVTVYVIGIRAPSEVLEAALDGG SQMRIMSRYLAIPHRIPLVVLTIOSVMDKRVFFQIYVMNGGIAQNMYLVANYAQ	
Query Match	14.0%	Score 84;	DB 1; Length 35856;
Best Local Similarity	60.4%	Pred. No. 1,2e-09;	
Matches 157;	Conservative 0;	Mismatches 100;	Indels 3; Gaps 1;
QY	210	tgaactggttcgaaagatcgccgcagaagtccaactctgcaagtgctggcaggagctgc	269
Dd	8478	TGAAGCTGTGGCGCCGCCGCGCAGACC GCCGCGAGGTGACCCGCAAGACGCCGAGCGCGCT	8419
QY	270	tgttaacgccttcacaagatgtgttctgtcga---ggctaigaagtcgcggcagaagcgctga	326
Dd	8418	GCTGGCGCGCTTTGGCCGAGGTTGGCGGACATGTCCTCCAAGGGCGACGAGAAGTGTCAC	8359
QY	327	gtcacggccttgtttcttcgcgtgaagcgcttaagcgccggctgcgacacggcgcgaacc	386
Dd	8358	CATCCCCCGCTTCTTGACTTTTCGAGCCGACACCCACCGTCCGCTCCACGCCGCCGAAACC	8299
QY	387	ggcgacctggcagcagaattgatcatctcggtctcaagcgcttgatctccgtgtgctc	446
Dd	8298	GCAAGACGGCGCACGCCGACGATTCGCGCGCGCTACAGCGTCAAGATGTCCTCGCGGCTC	8239
QY	447	cctgctgaagaagagcgctca	466
Dd	8238	CAAGCTCAAGCAAGACCCGCCA	8219

[illegible]

	gene				VVVAEYVEGKACGLLIYAERKKGAASA"
		7971 .8591	/gene-"CC1962"		
CDS		7971 .8591	/gene-"CC1962"		
			/note-"identified by match to protein family HMM"		
			/codon_start-1		
			/transl_table=1		
			/product-"c1ca protein"		
			/protein_id-"AAK3937.1"		
			/db_xref="GI:13423423"		
			/translation-"MTGEMGHEPLLVAFDEDTLTVDSPNAFLKWRAPGPRMSGVLR		
			LFTPLLIVFDNRNKGKLKAARVOELGATVAQIENRARAFAEAPEALRPDVAAYWA		
			RGMRAKAKMYNIWASPDILYAPFARGLGADLLIGTLRCDDRIILGGDGNCRKAK		
			EKVRLKEVRDPDYRLTAITDITSGDTEMLAIDAEKGRIFRGPA-		
	gene		complement(8670 .9302)		
			/gene-"CC1963"		
CDS			complement(8670 .9302)		
			/gene-"CC1963"		
			/note-"identified by match to TIGR protein family HMM		
			TIGR00493"		
			/codon_start=1		
			/transl_table=1		
			/product-"ATP-dependent Clp protease, proteolytic subunit"		
			/protein_id-"AAK3938.1"		
			/db_xref="GI:13423424"		
			/translation-"MWTDPVSAMNLVPWVEOTSRRERAEDIFSRLKERIFLNGP		
			VEDMASILICAOQLFLESENKKEIKIAYINSPGVAVAGLAITYDMOYIKSPYSTVM		
			GMASSGLSLAAGAAGORISILPNARIWHPGSGFGQASDIERRHEDIIKTRRRLN		
			EIVYHCGRTRYEEERTLDRDHEMSADEAKMWGLVDHYVSDRNAEAGAE"		
	gene		complement(9480 .10838)		
			/gene-"CC1964"		
CDS			complement(9480 .10838)		
			/gene-"CC1964"		
			/note-"identified by match to TIGR protein family HMM		
			TIGR00115"		
			/codon_start=1		
			/transl_table=1		
			/product-"trigzer factor"		
			/protein_id-"AAK3939.1"		
			/db_xref="GI:13423425"		
			/translation-"MOIIEKSGEGLSIRVFYGVASELATRTLEARIAEVAPOMNVGE		
			PGCVPTAHVRLRYTGKALMGVEITGOALNTTYYKLLENKILAPACOPILNSDDMDYIV		
			AGGEDLSFDAVEYMPREPERIDPISIELKPYRVSPDEVOALDELAKQARTIEPTT		
			GKSLKARDQDLIDVGTIDGVEFAGKKEAGALVLGSGFITFEEDULGAAPGDGD		
			VVVVFPPEYOARDLAGKDAEFATKQVEVARAPVGADDELARLGISDLAALTILLI		
			KSNLAGRYDNSSRFKRLRLALDVTDTHDFLPRLPVDAEPAGIMQOVENDKARGLIP		
			PENDETEDOLDKEYEKRIIERRVRLGLVLAIEGRKNVYVVIDELTPVAINBEARQVGA		
			OAOQVPMYRQRADLQAALARPIYEDKRVLLIFOKATIEKYSKDLEBDLPBET		
			GG"		
BASE COUNT		1854 a 3619 c 3396 g 2091 t			
ORIGIN					
Query Match		13.7% Score 82.4 DB 1 Length 10960;			
Best Local Similarity		56.7% Pred. No. 3.2e-09;			
Matches 152; Conservative 0; Mismatches 116; Indels 0; Gaps 0;					
OY	201 caacagctcactacgcttcgaagaatgccccagaagtccaacctgcaccagaagctcagc	260			
Dd	3382 CACAATAACCGCAGCTCGTTACGCCGATGCAGAGAGGCGGCAATCAACAGAACCGAC	3323			
OY	261 cgaagctcglttaagccttcccaagatgtgttgctcgtagagctatgaagctccgcgaag	320			
Dd	3322 CAAGGACCGCGCTGGAAGCGCTTATTGAAAGCGGTACTGCACTCGTGAATCGGCGCAAGA	3263			
OY	321 cctgaagctaacggccctgtttctccgctgagcggtcacaaggccccgctcgaccggccg	380			
Dd	3262 CGTGGCTTGTTCGGCTTTGGCACCTTCAGAGCGCGTGAACCGCGGCTGGAACCGCGC	3203			
OY	381 caaccgcgcacatcgcgagcagattgacatctcggcttcactcagcgcttgtcatctcgc	440			
Dd	3202 TAATCCCCTACCGGTAGAGACCGCTAACCGTCCCGCTCCCAAGAGCGCCCGCTTCCAGCT	3143			

```

OY      441  ttgctccctgctgaagaagcgctcacc 468
|||      |||||      |||      +
Db      3142  CGCGGAAGGTCTGAAGTCTTCGCTCACC 3115

RESULT  6
AF110185
LOCUS   AF110185
DEFINITION
Burkholderia pseudomallei strain 1026b Dbhb (dbhb), general
secretory pathway protein D (gsdp), general secretory pathway
protein E (gspe), general secretory pathway protein F (gspe), Gsps
(gspc), general secretory pathway protein G (gspe), general
secretory pathway protein H (gsph), general secretory pathway
protein I (gspl), general secretory pathway protein J (gspl),
general secretory pathway protein K (gspe), general secretory
pathway protein L (gspl), general secretory pathway protein M
(gspm), and general secretory pathway protein N (gspe) genes,
complete cds; and unknown genes.
AF110185
AF110185.1  GI:4139231

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Burkholderia pseudomallei.
Burkholderia pseudomallei
Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
Burkholderia; pseudomallei group.
1 (bases 1 to 20302)
Deshazer,D., Brett,P.J., Burnick,M.N. and Woods,D.E.
Molecular characterization of genetic loci required for secretion
of exoproducts in Burkholderia pseudomallei
J. Bacteriol. 181 (15), 4661-4664 (1999)
99350433
2 (bases 1 to 20302)
Deshazer,D., Brett,P.J., Burnick,M.N. and Woods,D.E.
Direct Submission
Submitted (30-NOV-1998) Microbiology and Infectious Diseases, T2N
University of Calgary, 3330 Hospital Dr. NW, Calgary, Alberta, T2N
4N1, Canada

FEATURES
source
Location/Qualifiers
1. 20302
/organism="Burkholderia pseudomallei"
/strain="1026b"
/db_xref="taxon:28450"
complement(<1..458)
/note="orfA: similar to Escherichia coli Oxf 688 and
Haemophilus influenzae gene H1154"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAD05169.1"
/db_xref="GI:4139232"
/translation="MSRARSGRNRLRGLAARAFRWILPTPGGTSRRFAAFS
PSPFLPMRDRIYPAITVEREDGCVVSGDIYHSAGLADANHYFINGNLPERW
RHRAPPIETGTGTGCTGCTCNFLATYAAARADSHCEHLFVSVEKHPAREDL"
697.. 975
/gene="dbhb"
/gene="dbhb"
697.. 975
/gene="dbhb"
/codon_start=1
/transl_table=1
/product="Dbhb"
/protein_id="AAD05170.1"
/db_xref="GI:4139233"
/translation="MNKOELIDAVVAQTGASKAQCTETDILLLEVIKAVSKDSVOL
IGGSGRSGARRATGRTGTGETIKIPAKYVKTAGKAFKDAVNR"
complement(1078..2421)
/note="orfB"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAD05171.1"
/db_xref="GI:4139234"
/translation="MNDPLEVYVLSGFLGAGKTTLLNHLIINARAGMKVAVIYVDLAA

```


Db	760	GGCGAAGCGTGGACACAGCTCTCTGAAAGTATCAAGAAGCGTGTGTGCAAAAGCGACTCA	819
Qy	322	ctgaagctacacggcgtctgtcttcctcgctgagcgcgtcaagcgcgcgtcgcacccggcgc	381
Db	820	GTTTCACCTGATCGGCTTCGGCAGCTTCGGCTCGGGCAAGCGCCACCGCGCACCGGCCCC	879
Qy	382	aaccgcgcactggcgcgcagatgacattccgcgcgtcctcctcagcgcgttgcgtctccgct	441
Db	880	AAACCCCAAGACGGCGGACCATCAAGATCCCGGCAAGCAGCGGTCAAGTTTCACGGCT	939
Qy	442	ggtccctcgtcgaagaagcgcgtaccgaagt	473
Db	940	GGCAAGCGCTTCAAGGACGACGTGACACGCG	971
RESULT	7		
LOCUS	AE004947/c		
DEFINITION	AE004947	11691 bp	DNA
ACCESSION	AE004947		
VERSION	AE004947.1		
KEYWORDS	GI:9951665		
SOURCE	Pseudomonas aeruginosa.		
ORGANISM	Pseudomonas aeruginosa Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; Pseudomonas.		
REFERENCE	1 (bases 1 to 11691)		
AUTHORS	Stover,C.K., Pham,X.O., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S., Hufnagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltzy,L., Tolentino,E., Westbrook-Wadman,S., Yuan,Y., Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Laidig,K., Lim,R., Smith,K., Spencer,D., Wong,G.K., Wu,Z., and Paulsen,I.T.		
TITLE	Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen		
JOURNAL	Nature	406 (6799)	959-964 (2000)
MEDLINE	20437337		
REFERENCE	2 (bases 1 to 11691)		
AUTHORS	Stover,C.K., Pham,X.O.T., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S.L., Hufnagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltzy,L., Tolentino,E., Westbrook-Wadman,S., Yuan,Y., Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Laidig,K., Lim,R.M., Smith,K.A., Spencer,D.H., Wong,G.K.-S., Wu,Z., Paulsen,I.T., Reizer,J., Sailer,M.H., Hancock,R.E.W., Lory,S. and Olson,M.V.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-MAY-2000)		
FEATURES	Department of Medicine and Genetics, University of Washington Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA		
source	location/Qualifiers		
gene	1. .11691		
CDS	/organism="Pseudomonas aeruginosa"		
gene	/strain="PA01"		
CDS	/db_xref="taxon:287"		
gene	76. .1485		
CDS	/gene="PA5346"		
gene	76. .1485		
CDS	/gene="PA5346"		
gene	/codon_start=1		
CDS	/transl_table=11		
gene	/product="hypothetical protein"		
CDS	/protein_id="AG08731.1"		
gene	/db_xref="GI:9951666"		
CDS	/translation="MTEALASADAPARHOVILQILAKIGTACREVSDSAELPARRYQAVLIDAVGTLIVLTPQSOGLDLARLTLETKLTAIPERLERMKGHEHLRIPLPITSSPCLYDERLQEPRLLESQGPILLETLASDNERGLGKASAFQEPLENILNLRDDDDPALEISQVQAFTRARIQOREQTEIEPLPEQAKIITKLRVDPNSVDITGLVETDPALAAQVYVMAASPYAAPGKIRTSDEAIYVGLGDLVIMLALGKSLSPKQPOAPPYMOOAITATVAVIEGLRAMPRELPRPSGLSYGLGLHNGYVLAHPPHPSICRHLTEVQVPHGLVATIEOHLITRIGAMLMRWMDMDLIDYCALRQODPSYTGPNNAVYANILCTNRLIRNLSIGDGIQOAIIPALYRLITTEBKADAVKYLEAEALRELAQAFNRKPG"		
gene	complement(1557..1949)		
CDS	/gene="PA5347"		
gene	complement(1557..1949)		
CDS	/gene="PA5347"		
gene	/codon_start=1		
CDS	/transl_table=11		
gene	/product="hypothetical protein"		
CDS	/protein_id="AG08732.1"		
gene	/db_xref="GI:9951667"		
CDS	/translation="MKFPLIMLMGRLMAKSRONPAFQKLEKDLVYOLHLLDGVANHVLVRLDRSSKRGTHLPQAFSLGFDGVAIGFAAMTSKNPOLAFMQGIQNKDIOIGNPALVWFGVLGYLYLKKPAEKKA"		
gene	complement(2085..2357)		
CDS	/gene="PA5348"		
gene	complement(2085..2357)		
CDS	/gene="PA5348"		
gene	/codon_start=1		
CDS	/transl_table=11		
gene	/product="probable DNA-binding protein"		
CDS	/protein_id="AG08733.1"		
gene	/db_xref="GI:9951668"		
CDS	/translation="MRKPELAALAEKADLTKEDQANRYLNLDELITGALNRKQSVTLVGEFPIORHRGAFRGKMPQGPVKIKASNTVAFKPKALRDVNV"		
gene	complement(2559..3713)		
CDS	/gene="PA5349"		
gene	complement(2559..3713)		
CDS	/gene="PA5349"		
gene	/codon_start=1		
CDS	/transl_table=11		
gene	/product="probable rubredoxin reductase"		
CDS	/protein_id="AG08734.1"		
gene	/db_xref="GI:9951669"		
CDS	/translation="MSERAPVITGTGLAGYNLARERKKLDGETPLMITRNDGRSISKPLSTGEKKNADAPGLMAEPAEOLNRIILTHRTVGTIDGHORIMVGESEVYRLVLVAMGAEPIRYVPEDDADALYPIINDLEDVAFRQAPAGKRRVLLAGLIGCEFRDLVSSGGQOLDVVAPEQGVMPGLHPAAKVAQAGLEGVLRHLGPIVLAISKABEGLAEHLSDSEVLPCLDVSAVAGRLRLAEAGLAIVRGIVYDRSLTSHANTYALGADAEVYDNLILVYPLMCAKARLAOTLEAGNSOYVAGMPVTVTPACPIVSPPRGMDGCVLVSGSGRDTLVLCRDNRARVIGALVGAIVNEMLAINKELPGIMN"		
gene	complement(3765..3932)		
CDS	/gene="PA5350"		
gene	complement(3765..3932)		
CDS	/gene="PA5350"		
gene	/codon_start=1		
CDS	/transl_table=11		
gene	/product="rubredoxin"		
CDS	/protein_id="AG08735.1"		
gene	/db_xref="GI:9951670"		
CDS	/translation="MRKMQCVVCGFIYDEALGLPEGIPACTRWEDIPADWVCPDCGVGRIDFEMIRIA"		
gene	complement(4116..4283)		
CDS	/gene="PA5351"		
gene	complement(4116..4283)		
CDS	/gene="PA5351"		
gene	/codon_start=1		
CDS	/transl_table=11		
gene	/product="rubredoxin"		
CDS	/protein_id="AG08736.1"		
gene	/db_xref="GI:9951671"		
CDS	/translation="MKRWQVVCGLIYDAKGPBEGTEAGTRWEDVPEDMLCPCDGVGRIDFEMIRIG"		
gene	complement(4417..4818)		
CDS	/gene="PA5352"		
gene	complement(4417..4818)		
CDS	/gene="PA5352"		
gene	/		

gene
/gene="g1cf"
/note="PA5353"
complement(4823..6049)
/gene="g1cf"
/codon_start=1
/transl_table=11
/product="glycolate oxidase subunit g1cf"
/protein_id="AG08738.1"
/db_xref="GI:9951673"
/translation="MOTNLCDSARQAEAEERILSRSCVHCFCFNATCPYQVEIGADE
LDDPRLRITVIAKQFLREDAESERTRLHLDRCISCRNCETTCPSGYEYKHLIDIGDIL
LADLPKGFRRRLRDLRLRLAPRPGFGLAAGRLRILPEAMOGGLPRVAAQGLR
RPPRRARRLILEGCGVOPALSPNTAAARVILRISVYMPAAAGCCGALDFPILDA
QOEGRARARNIDAMWPIRLREGAEALIVQAAASCGAFYVDYATLLDDPRYADKAEVA
AARLDIYEVIRDEPLESLGRLRCDQRLAFICPCTTGQAQRLGAEVAELRLGFLITEV
ADHCLICGAGCTYSLTGPAALISQRLRDNRLDALESGAPDVTYANVGCAHLDGARTE
VRRHIELLDQALPGGE"
complement(6059..7138)
/gene="g1cf"
/note="PA5354"
complement(6059..7138)
/gene="g1cf"
/codon_start=1
/transl_table=11
/product="glycolate oxidase subunit g1cf"
/protein_id="AG08739.1"
/db_xref="GI:9951674"
/translation="NSQPYFDDPAGADHAGLLEERYDALRRAPRLRVGDSKAFIC
RPVAGPLDLAHRGIVATDPCELVITTAAGTPIDELQAVTLARGQMILACAPGFGG
ATVGGVLASGLSGRRPMPAGAVRDLVLTGRLISGLGTHLRFGGEMKNVAGYDLSRLIL
ASFGCLGVATEVSEFVLPMPRQAGLSTLDLFLVAMALKLAEWGGQPLPISACHDGG
ALRLRLEGGGVSRSARLDRLGGEALDDGMADLRQORLPFPDDPRPLMRSLPLAAPE
TLPGLGGLDGGGALRMLKSDAPADINAMTAHGHMTCYRAGHLDAIRPPLSPGLLIL
GHRALKRQLDLPQGIFFNPGRMAYEL"
complement(7138..8637)
/gene="g1cd"
/note="PA5355"
complement(7138..8637)
/gene="g1cd"
/codon_start=1
/transl_table=11
/product="glycolate oxidase subunit g1cd"
/protein_id="AG08740.1"
/db_xref="GI:9951675"
/translation="NNILYDERLDPLPQVNDKGILAERLRLPDLELLHAEDLRPY
EDGDSAYTCPLVLALPERIPIROVGGVILALICHRILKVPVARGAGTGISGALPLENVL
LLVMAFRRIEIDPLGRPARPQGVRLILASOAAAPRGAYPADPSSQINCSIGNV
ANAGVCHCKRYGLTVNNHLLOYDITLTGERSLGSASALDSAGFDLALTLFGSBSGLIG
VVEYTVRLRLRPVAKVLASFDVESAGRAVADLIGAGIVPAGLMMNMLSTRAE
DTHAGYVPDAALAILCELDEVEADVEDHCERVELPEAGATSVROAOEAEROFW
AGRNKAFLPANGELSEPRAYCDMSGSLPRRLPVHLVIAELSTETGLRVANFYHAGDGM
HLLILFANLNGELSERALGRIILELCVAVGGSLTGHGVRKINOMCAQFNADEL
TLFHVAKAFDPAGLNLPGKNVPLTLRCAERGAHHVHGRLPPELELRP"
8842..9597
/gene="g1cc"
/note="PA5356"
8842..9597
/gene="g1cc"
/codon_start=1
/transl_table=11
/product="transcriptional regulator g1cc"
/protein_id="AG08741.1"
/db_xref="GI:9951676"
/translation="MOKROVADOVAERIERLIVDVLKVGQALPSESRVLVAKIGCSRS
ALRELRALNRGIIIDLEHGRGSFVADLRNADVSPLMHFGSOPRTLYLTLVRL
EGEARLALRGTEADPVLRLARREEMLASHEETOPIDPREHARDDHAFRLALSPSH
NPVLTLOSINELLSTVFASVNNVTHRPQKROIDRQARLRYAALREQPODAQRA
AREHHSIRDLNRLTEDEQRILVATYTLRNLWG"
9677..10213
/gene="PA5357"
9677..10213

Query Match	11.4%	Score 68.2	DB 1	Length 393
Best Local Similarity	53.1%	Pred. NO. 9.3e-06		
Matches 145	Conservative 0	Mismatches 134	Indels 0	Gaps 0

Query	202	accagtcgtcctggtttcgaagatcgcccaagaatccaaactgacaaagcgtcagccag	261
Db	2251	AAACGAACTAGCCGCGCGTATCGCGGAAAGAAAGCCGATCTACCAAGAACAGGCGCAAT	2292
Qy	265	gctgctgtaagcgtcttccagatgtgttcgtcgaagcatatgaagtcggaagcgtg	324
Db	2291	CGGTTCTCAACCCCTGTTGGATGAATCACCGGCGCGCTGAACCGCAAGGACACG	2232
Qy	325	aagtcacccgacctgtcttcgcgtgagcgcgtaagcccggtctgcacggcgcaac	384
Db	2231	ACCTGTGCTCGTTTCGGACCTTCTCGCAACGCGCATTCGGAGACCCGACGGGGAAG	2172
Qy	385	ccgagcactgagcagacagattgacattccggtcttctaagcgcttgatctcgtg	444
Db	2171	CGCAGACCGGCGCCAGCGGGAAGATCAAGCGCAACACCGCTCCCTTCAAGCCGGC	2112
Qy	445	tccctgctgaagaagcgcgtacacgagtgactctgtctgtag	488
Db	2111	AAGGCCCTGCGCGACGCGGTCAACTGACCGCGAGAGCCCGAG	2068

RESULT	8
LOCUS	AF090448 393 bp DNA BCF 04-OCT-1998
DEFINITION	Rhizobium leguminosarum Hu-1-like protein gene, complete cds.
ACCESSION	AF090448
VERSION	AF090448.1 GI:3694945
KEYWORDS	
SOURCE	Rhizobium leguminosarum.
ORGANISM	Rhizobium leguminosarum Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Rhizobiaceae; Rhizobium.
REFERENCE	1 (bases 1 to 393) Liu,S.T., Chang,W.Z., Cao,H.M., Hu,H.L., Chen,Z.H., Ni,F.D., Lu,H.F. and Hong,G.F. A HU-1-like protein binds to specific sites within nod promoters of Rhizobium leguminosarum J. Biol. Chem. 273 (32), 20568-20574 (1998)
TITLE	2 (bases 1 to 393)
JOURNAL	Hu,H.L. and Hong,G.F. Direct Submission
MEDLINE	Submitted (05-SEP-1998) Shanghai Institute of Biochemistry, Chinese Academy of Sciences, 320 Yueyang Road, Shanghai 200031, China
REFERENCE	
FEATURES	Location/Qualifiers
SOURCE	1..393
gene	/organism="Rhizobium leguminosarum"
gene	/strain="8401"
gene	/db_xref="taxon:384"
gene	1..393
gene	/gene="hurle"
gene	52..327
gene	/gene="hurle"
gene	/note="hr1"
gene	/codon_start=1
gene	/transl_table=11
gene	/product="HU-1-like protein"
gene	/protein_id="AAC62499.1"
gene	/db_xref="GI:3694945"
gene	/translation="MNKELVSAVAEKAQGLTKSDASAVDAVDYQAELEKNGDITL AGRFSYVSHRAAKGNPNSTGAAYDIPARNVPRFTPGKGLDAVNC"
BASE COUNT	91 a 101 c 115 g 86 t
ORIGIN	

gene 2945. .3220
/gene="AGR_C_2331"
2945. .3220
/gene="AGR_C_2331"
/note="DNA-BINDING PROTEIN HRL18"
/codon_start=1
/transl_table=11
/product="AGR_C_2331p"
/protein_id="AAK87057.1"
/db_xref="GI:15156311"
/translation="MNKNELVSAVAERAGLTKADAASAVDAVFETVQGEELKNGDIDL
AGGFSFSRRERDASGRMRPSTGAEEVDIPARNVPRFSAGKGLDAVNS"
3234. .3398
/gene="AGR_C_2333"
3234. .3398
/gene="AGR_C_2333"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="AGR_C_2333p"
/protein_id="AAK87038.1"
/db_xref="GI:15156312"
/translation="MKCPAAGMRLPQAAHEETCFDALMPLSGVRMPSKALSGGPVSPILM
VWIESLRAPV"
3537. .5762
/gene="AGR_C_2334"
3537. .5762
/gene="AGR_C_2334"
/note="(AL117669) putative secreted protein"
/codon_start=1
/transl_table=11
/product="AGR_C_2334p"
/protein_id="AAK87059.1"
/db_xref="GI:15156313"
/translation="MTLDPRGIRKMAFRFESHVALTAALITSAAPFAAEPEVFNRIASPP
VAKNLPADQADKSVTSAEIVTASDEGKTLIYSDSLPAGIGFDITDDAKAPAGACALLIL
OGEPISVAAGAKAGIAGVNTSENRNAPKPGKLTATYADIAATKITEASCDLGGDDVSAVS
DKTFAIATENRDEEVNEGALPQMPASGLTYVGLKQDAADCATLTKIDTGLSEISRP
EDPEEFPAISANGELIALTMOENNHIVIIIDKRTGIVKTHFSAGTVNLGDIOTKIDGDL
FTTQDEQRKREPPDAVKMLDNDRIIVIANEGMOMGSGRGTIFEDTGTKLLEAGASLREH
AVAAIGHGPEKRSKAKGIEPEGLEAATPDGQKFFVLAEASVVAAYKDTCAEQLQ
ILPSTISPEGVAIIPSRNLITANEDADLAEGGARSHMLYERAEQAAAYPOLVASVIV
DGAPIGKMLASGLIYADARKPGIILYAVSDSYVGNQPSITYTIDATKKPAATISVLEPKVD
GIAAKKILIDEGITLDDKGGFWLASBDSAKILPHALYNVAKKEITKAEIALPKRLNG
DTRFEFEEVITWAGKDDATLMAAVQREMGTDKEKGFVKLYSNPKSKMGAVRPLDMLR
EAGWVGESEITIAQGDNNVIVERNDOIGDKARKIKLRYAILEDLPKGIIGGLPLVAKO
EAHDELPLDKAQGTNGVVDKLEGFTEFKAGTAAYVTDNDGVDSSEGTLPFTVNLKAL
N"

gene complement(5755. .5874)
/gene="AGR_C_2336"
complement(5755. .5874)
/gene="AGR_C_2336"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="AGR_C_2336p"
/protein_id="AAK87060.1"
/db_xref="GI:15156314"
/translation="MPRHASCKIGDTAHGSLGDTKKKPGYARLRFSPVPLGTS"
complement(5968. .6105)
/gene="AGR_C_2337"
complement(5968. .6105)
/gene="AGR_C_2337"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="AGR_C_2337p"
/protein_id="AAK87061.1"
/db_xref="GI:15156315"
/translation="MDEFMSIDAFTYTLALILPVIIFPRRWQDSEKEDPDKSSSPG
R"
gene complement(6407. .6862)
/gene="AGR_C_2338"

complement (6407..6862)
/gene="AGR_C.2338"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="AGR_C.2338p"
/protein_id="AAK87062.1"
/db_xref="GI:15156316"
/translation="MSRRFDKRISEFIIIGNNRTFKHDCCKAPFALHNENKLV
HSRRPAQVYKAVRASPRDPKAPVAGFLPLAFRPGKSCPEKQNFKALAVFEE
RSATYATADAMIRKTRRAISSVGNAPRLHGRCEFEESLIHHFLKTS"
7577..7942
/gene="AGR_C.2340"
7577..7942
/gene="AGR_C.2340"
7577..7942
/gene="AGR_C.2340"
/note="NADH DEHYDROGENASE I CHAIN A 1 (NADH-UBIQUINONE
OXIDOREDUCTASE CHAIN A 1)"
/codon_start=1
/transl_table=11
/product="AGR_C.2340p"
/protein_id="AAK87063.1"
/db_xref="GI:15156317"
/translation="MTELLSYVYIPIAFIPIGIALVIGLALLIAPAVAAKADPEKLSN
PFCGNAPADAMKPDIREPLVSLIFPIIDLEVAFLEFPWAVSFGEMGFWGSMVFL
AVLTIGFTIEKKKALENA"
7912..8517
/gene="AGR_C.2341"
7912..8517
/gene="AGR_C.2341"
/note="NADH DEHYDROGENASE I CHAIN B 1 (NADH-UBIQUINONE
OXIDOREDUCTASE CHAIN B 1)"
/codon_start=1
/transl_table=11
/product="AGR_C.2341p"
/protein_id="AAK87064.1"
/db_xref="GI:15156318"
/translation="MEESGACGAGMGNNSSTLVAPQPKIIDPATGKPYGSDAYFTEI
NDELAKGEIYSTDELITWARTGSLMMQFGLACAVEGIMASGPRYIDMERGVAP
RASPSQSDVMVAGTILTNKMAPALKRYDQMEPRRYVISMGSANGGYYHYSYAAVVR
GCDRVVPIAVPYGPGPTPAEALLYGVLLIQRKIRRTGIYR"
8539..9141
/gene="AGR_C.2342"
8539..9141
/gene="AGR_C.2342"
8539..9141
/gene="AGR_C.2342"
/note="(A)245398) nuoc1"
/codon_start=1
/transl_table=11
/product="AGR_C.2342p"
/protein_id="AAK87065.1"
/db_xref="GI:15156319"
/translation="MSEALNDLILAVYKEARGLSVASDIAYGELTNTPEENYALLT
FLRDVQGVGFNIIDICVDMPQREKREDVYHLISPRONLRVGIKLOVADEGVSS
TPVYQAGFMEAEAMDYGIPEGHKDIRLITDYGEEGHRLKDPFVYGVFEVRIDD
VLRVLVYPVVLKQEFNFDLSFREGTEIYLPDEKAKQ"
9138..9566
/gene="AGR_C.2344"
9138..9566
/gene="AGR_C.2344"
/note="hypothetical protein SPAC29B12.13 - fission yeast
(Schizosaccharomyces pombe)"
/codon_start=1
/transl_table=11
/product="AGR_C.2344p"
/protein_id="AAK87066.1"
/db_xref="GI:15156320"
/translation="MSEAPAREHGQCICGAVGATGAGREYGCISMCRRNSGAF
MAVECTIDENEDRLGYTSSKMGECFCNCGSTVSGSKGKHPAVSLQFDRPS
SFTVYQIIFTEDEKRSYISFAQTONTMGPEFIAMITSAEH"
9568..10758
/gene="AGR_C.2346"
9568..10758
/gene="AGR_C.2346"
/note="NADH DEHYDROGENASE I CHAIN D 1 (NADH-UBIQUINONE

[illegible]

BASE COUNT	214 a	319 c	331 g	106 t
Query Match	11.18;	Score 66.8;	DB 1;	Length 970;
Best Local Similarity	57.08;	Pred. No. 1.8e-05;		
Matches 122;	Conservative 0;	Mismatches 92;	Indels 0;	Gaps 0;
OY	256	caggccgagctgctgttaacgctccacagatgtgttcgtcagagcatgaatcgcgc	315	
DB	326	CAGGCGCCGCGAGCGTGTGACGCGCGGTCTCTGACGCGCTTCGTCGCGTCGCGGCG	385	
OY	316	gaagagccgaagctccacgagctgttcctccgttagagcggtcaagcgccgagctgcacc	375	
DB	386	GACCGGGTCTGCGTACCGGCTTCGCTTCCTTCGAGAAAGGTGACCGCGCCGCTTAC	445	
OY	376	gagccgacccgcgcgacgcgcagagatgtacatccgcttccttaagcggttcgcatc	435	
DB	446	GCCGCGACCCGCGAGAGCGGGGAGCGGGGTGCGGTCAGAAAGACCTCGTGCAGCGCTTC	505	
OY	436	tcgcgtgctccctgctgctgaagaagcgctacgcg	469	
DB	506	CGCGCGGGCGCAGGGGCTTCAAGAGACTGTGTGACCG	539	
RESULT 11				
SCIC2				
LOCUS	42210 bp	DNA	BCT	15-JAN-1999
DEFINITION	Streptomyces coelicolor cosmid 1c2.			
ACCESSION	AL031124			
VERSION	AL031124.1 GI:3355667			
KEYWORDS	3-isopropylmalate dehydratase large subunit; 3-isopropylmalate dehydratase small subunit; 3-isopropylmalate dehydrogenase; branched-chain amino acid aminotransferase; carbonyl transferase; delta-1-pyrroline-5-carboxylate dehydrogenase; glx; glutamyl-L-tryptophan synthetase; histone-like DNA binding protein; hydrolyase; llye; leuB; leuC; leuD; lyase; secreted lyase; transfer-RNA-Gln; transfer-RNA-Glu; ureAB; urease alpha subunit; urease beta and gamma subunits; ureC.			
SOURCE	Streptomyces coelicolor A3(2).			
ORGANISM	Streptomyces coelicolor A3(2)			
REFERENCE	Bacteria: Firmicutes: Actinobacteria: Actinomycetidae; Actinomycetales: Streptomycinae; Streptomycetaceae; Streptomyces.			
AUTHORS	1 (bases 1 to 42210)			
JOURNAL	Murphy, L. and Harris, D.			
REFERENCE	2 (bases 1 to 42210)			
AUTHORS	Parkhill, J., Barrell, B.G. and Rastandream, M.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (23-JUL-1998) Streptomyces coelicolor sequencing project			
REFERENCE	Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridg			
AUTHORS	CB10 ISA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.			
TITLE	David A. Hopwood, [3] John Innes Centre, Norwich Research Park,			
JOURNAL	Colney, Norwich, Norfolk NR4 7UH, UK			
REFERENCE	3 (bases 1 to 42210)			
AUTHORS	Redenbach, M., Kleser, H.M., Denapate, D., Eichner, A., Cullum, J.,			
TITLE	Kinash, H. and Hopwood, D.A.			
JOURNAL	A set of ordered cosmids and a detailed genetic and physical map			
REFERENCE	for the 8 Mb Streptomyces coelicolor A3(2) chromosome			
AUTHORS	Mol. Microbiol. 21 (1), 77-96 (1996)			
TITLE	97000351			
JOURNAL				
MEDLINE				
COMMENT				

FEATURES

SOURCE

Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Troph et al., Nucleic Acids Research, 22(12):4768-4779(1994)) and the Frameplot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.go.jp/jun/cgi-bin/frameplot.pl>. **CAUTION:** We may not have predicted the correct initiation codon. Where possible we choose an upstream codon (atg, gtg, ttg or att) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring subclones. Cosmid IC2 lies between BD9 and 7A1 on the AseI-B genomic restriction fragment.

Location/Qualifiers

1..42210

location="Stratocarnegie CGe1[Col]or AT(2)"

```

gene
CDS
1..1228
/gene="SC1C2.01"
<1..1228
/gene="SC1C2.01"
/note="SC1C2.01, probable delta-1-pyrroline-5-carboxylate
dehydrogenase, partial CDS, len >408 aa; similar to many
e.g. TR:050443 (EMBL:AL010186)
delta-1-pyrroline-5-carboxylate dehydrogenase (M.
tuberculosis) (543 aa), fasta scores; opt: 1753 z-score:
2214.2 E(1): 0, 65.1% identity in 407 aa overlap, and
PF072_HUMAN delta-1-pyrroline-5-carboxylate dehydrogenase
(563 aa), fasta scores; opt: 1273 z-score: 1730.4 E(1): 0,
48.3% identity in 408 aa overlap. Contains PS00687 and
PS00070 Aldehyde dehydrogenases glutamic acid active site
and cysteine active site and Pfam match to entry PF00171
aldehyd, Aldehyde dehydrogenases, score 114.70, E-value
1.8e-30"
/codon_start=2
/transl_table=1
/product="delta-1-pyrroline-5-carboxylate dehydrogenase"
/protein_id="CAI19968.1"
/db_xref="GI:3355668"
/db_xref="SPTREMBL:O86502"
/translacion="IDSECELIDFWRVNVHYARNILAEQPPANSGVYNNRMDHPLE
FYATLTPNFSIAIANLEPTALPMGNVYVYVMPSPQTAANVIAIMOLEEALPGVIV
LVTGCVIAIALEHRLDLAGIHETGSGTKPTFOHLMTKGNINEXRYRPLVGEENG
DPLVAFSPADRVAULTKATLRGAFFEQYQKCSATSAAYPAISMDGFEFEERAAEDVY
TMDVDTLISNFTGAVIDERFRAKNKAAIDRAKEDTCTIVAGSGIDSVCVFYRPTVY
ECTDENVEFRIEFGPLAVHYDSDADDAIDAMLTMESVSDIALTGVSISNDRA
AAYTEKRIYVAGNFFYINDKSTGAVVGOQPGGGRASTNDKAGAPQMLMTWTLTRAA
KETLVAIPDYRPHMG"
1..3807
/note="true overlap with cosmid 809"
2..1189
/gene="SC1C2.01"
/note="Pfam match to entry PF00171 aldehyd, Aldehyde
dehydrogenases, score 114.70, E-value 1.8e-30"
464..487
/gene="SC1C2.01"
/note="PS00687 Aldehyde dehydrogenases glutamic acid
active site"
548..583
/gene="SC1C2.01"
/note="PS00070 Aldehyde dehydrogenases cysteine active
site"
1246..1283
/note="hairpin_loop with 18bp stem"
misc_feature
misc_feature
misc_feature
misc_feature
misc_loop
stem_loop

```

```

RBS
1522. .1527
/Note="possible RBS upstream of SCIC2.02"
1537. .1668
/gene="SCIC2.02"
1537. .1668
/gene="SCIC2.02"
/Note="SCIC2.02, questionable ORF, len: 43 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein SCIC2.02"
/protein_id="CA19971.1"
/db_xref="GI:3355671"
/translation="MAETTVRRVRLSRTSDSRKNAALQSLDRDNGATGH"
1744. .2787
/gene="leub"
1744. .2787
/gene="leub"
/Note="leub"
/Note="SCIC2.03, leub, probable 3-isopropylmalate
dehydrogenase, len: 347 aa; similar to e.g. LEU3 CORGL
3-isopropylmalate dehydrogenase (EC 1.1.1.85) (340 aa),
fasta scores: opt: 1450 z-score: 1318.1 E(): 0, 68.0%
identity in 338 aa overlap. Contains PS00470 isocitrate
and isopropylmalate dehydrogenases signature and Pflam
match to entry PF00180 isodh, isocitrate and
isopropylmalate dehydrogenases, score 356.60, E-value
1.3e-105"
/codon_start=1
/transl_table=11
/product="3-isopropylmalate dehydrogenase"
/protein_id="CA19970.1"
/db_xref="GI:3355670"
/translation="MSRSLNAVIPGSGIGGVAAEGLKVLPODVKLETFEPD
GARYHATGERTLPADADALKADHAIILGAGDSVPGVLEKRLKRAEPHNVN
LRPKLPGVATPLAGPEIDFVYRRETSPTGNGTIRKGEHEVATSVNTAY
GVERVVDFAKRAQAPRKRLTLVKNVNLFAHLMTINRYAAEPEVTTYLHY
DAATFLVDPARPDVITDNLFGDIITDLAAVSGGIGVAGSININSGDFPFEP
VHGSAPIAGCGAKPTATVLSVALLHLGVEDEARIEDAVSADLGERDLPARST
SEIGDTLAARVAG"
1759. .2766
/gene="leub"
/Note="Pflam match to entry PF00180 isodh, isocitrate and
isopropylmalate dehydrogenases, score 356.60, E-value
1.3e-105"
2458. .2517
/gene="leub"
/Note="PS00470 isocitrate and isopropylmalate
dehydrogenases signature"
3038. .3042
/Note="possible RBS upstream of lve"
3052. .4140
/gene="lve"
3052. .4140
/gene="lve"
/Note="SCIC2.04, lve, probable branched-chain amino acid
aminotransferase, len: 362 aa; similar to many, e.g.
lve BACSU putative branched-chain amino acid
aminotransferase (362 aa), fasta scores: opt: 1047
z-score: 1401.5 E(): 0, 44.7% identity in 360 aa overlap.
Contains PS00770 Aminotransferases class-IV signature"
/codon_start=1
/transl_table=11
/product="branched-chain amino acid aminotransferase"
/protein_id="CA19971.1"
/db_xref="GI:3355671"
/translation="MTPTIELKPSAHLPSDESEKAILANGFGFHFTHDHWITKTE
GRCMHGOLVPAVSLIDPAMVLYHAGELFEGIKAVRPDGSVATRPKNGARQA
SSRIQMBELPVDTFACDAIVADKRWPARHGSESLIRPMITVGLGVRPN
EYLPYIYASPAAGATPFCGVRPVSITWSEDRVRAVGGMGDKTGGNTAAASILVAD
AKGCDVOCYLDALERKWEELGKMLTFYGNKIVTPTSLGSLIEGTVTRSLTLVARD
LGVEABGRVSVDMORDSENGLTEVFACGTAAVITPVGTVRKAGAQMOGSGGETGE

```

```

misc-feature
3763. .3852
/Note="PS00770 Aminotransferases class-IV signature"
gene
complement(4214..5644)
/gene="SCIC2.05c"
complement(4214..5644)
/gene="SCIC2.05c"
CDS
Query Match 11.1%; Score 66.8; DB 1; Length 42210;
Best Local Similarity 57.0%; Pred. No. 1.2e-05;
Matches 122; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 256 cagggcggagcgcgcgtgtaacgccttcacagatggttcgtcgaagcgtatgaatcgcgc 315
DB 41235 CAGGCGCGGACGCGTGTGACCGGGCTGCGACGCGCTGCGTGTGCGTGTGCGGCGC 41294
QY 316 gaagcctgaagcctacacgcgcgttcctccgctgagcgcgtcaagcgcgcgcgcgcac 375
DB 41295 GACCGGGTCTCGTGTGACCGCGTGTGCTTCTCGAGAGGTGACCGCGCGCGCTTAC 41354
QY 376 ggcgcgaacccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 435
DB 41355 GCCCGCAACCCGCAAGACGCGGCGGTGCGGTGCAAGAGACCTCGCGCGCGCTTC 41414
QY 436 tccgctgctccctgctgtaagaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 469
DB 41415 CGCGCGGCGCAAGGCGCTTCAAGGACCTGCGACCG 41448
RESULT 12
LOCUS AE005903 11684 bp DNA BCT 28-MAR-2001
DEFINITION Caulobacter crescentus section 229 of 359 of the complete genome.
ACCESSION AE005903 AE005673
VERSION AE005903.1 GI:13423852
KEYWORDS
SOURCE
ORGANISM
Caulobacter crescentus.
Caulobacter crescentus.
Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
Caulobacter.
1 (bases 1 to 11684)
Nierman,W.C., Feldblum,T.V., Laub,M.T., Paulsen,I.T., Nelson,K.E.,
Eisen,J., Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,J.R.,
Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B.,
Ely,B., Deboy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L.,
Hart,D.H., Kolonay,J.F., Smit,J., Craven,M., Khouli,H., Shetty,J.,
Berry,K., Uitterlinden,T., Tran,K., Wolf,A., Vamathevan,J.,
Ermolaeva,M., White,O., Salzberg,S.L., Venter,J.C., Shapiro,L. and
Fraser,C.M.
Complete genome sequence of Caulobacter crescentus
Proc. Natl. Acad. Sci. U.S.A. 98 (77), 4136-4141 (2001)
2 (bases 1 to 11684)
Nierman,W.C., Feldblum,T.V., Paulsen,I.T., Nelson,K.E., Eisen,J.,
Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,J.R., Potocka,I.,
Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B.,
Laub,M.T., Deboy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L.,
Hart,D.H., Kolonay,J.F., Smit,J., Craven,M., Khouli,H., Shetty,J.,
Berry,K., Uitterlinden,T., Tran,K., Wolf,A., Vamathevan,J.,
Ermolaeva,M., White,O., Salzberg,S.L., Shapiro,L., Venter,J.C. and
Fraser,C.M.
Direct Submission
Submitted (31-JUN-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
1. 11684
/organism="Caulobacter crescentus"
/db_xref="taxon:69394"
/complement(73..1236)
/gene="CC2328"
/complement(73..1236)
/gene="CC2328"
FEATURES
source
TITLE
JOURNAL
MEDLINE
AUTHORS
REFERENCE

```

```

/note="Identified by match to PPM protein family HMM
PF00441"
/gene
/codon_start-1
/transl_table-11
/product="acyl-CoA dehydrogenase family protein"
/protein_id="AAK24299.1"
/db_xref="GI:13423853"
/translation="MSGFWLIEQOEAIREGVANVCAAFDDEYMRIDETGEPEPAFVA
AIAEGWLVGAMPENVGAGLGLTEAAMQSAVSGAFSGASAIHLNIFGMPILVK
FGDEORNLRLPRIISGEDKCFAPVPPNSGLTSSLETREAEKVEGGYRLNGRIWYT
GABRANKIILITARTIPKOCAPKPHOGISTFYTBRENTIARKEPIPMGRKAYECMMLFE
DLVVPADLVGEERKQYILKILNPERVIFVPAEITCGAALAKATTYANERIVGR
PIQONQVAPHLAKSWAELEPANLAFKASLTDQGEKCGAENAAATYLGAEQFHC
EAVLVAGMGVYKEEDYERYFREAMTARIAPVSRREMIINFIAERYLGPKSY"
/complement(1265..2281)
/gene="CC2329"
/complement(1265..2281)
/note="Identified by match to PPM protein family HMM
PF01266"
/gene="CC2329"
/codon_start-1
/transl_table-11
/product="D-aminic acid oxidase family protein"
/protein_id="AAK24300.1"
/db_xref="GI:13423854"
/translation="MKLRGARIAVAGALGSAVALRLAQRGFRVTVDPSPGDANA
SSVAGMLAVSEALPDPASRNHLIDIKRARDLPDPAHYLGLEHGVRIKESDWM
RSVAVSLGALGEPAGAFIAEDWIDRAGILALRGASBAGVAFRLAVCGEAGVL
ELADGGEPPDITLIATPGILOGEITPERMISGILIRLAACDDAPAVYRGEG
VYIAPGVAVAGATMEGRDILPDELTATGLIRAAALRDLDLESATYEVGVRYT
PDGLPLGVMSRTPEVMLAVGARRNGWLFAPLVDMVAAYLFGDNLGPDALAMARRRE
KPYT"
/2495..2902
/gene="CC2330"
/2495..2902
/gene="CC2330"
/note="Identified by match to protein family HMM"
/codon_start-1
/transl_table-11
/product="transcriptional regulator, Cro/Ci family"
/protein_id="AAK24301.1"
/db_xref="GI:13423855"
/translation="MNDPRAEQAHFVDLYVGARLRRRKMGLSQTVADALGITTO
QIQIKERGANRISKLIDAKLQAPVSTFEBLEPDGVDGDFAGQRTAEVSTPE
GLELASLPRLADRLRRRVVDLVRAYVDADAP"
/3209..3487
/gene="CC2331"
/3209..3487
/gene="CC2331"
/note="Identified by match to protein family HMM"
/codon_start-1
/transl_table-11
/product="DNA-binding protein HU"
/protein_id="AAK24302.1"
/db_xref="GI:13423856"
/translation="MNVSDLYDAVAADKELTKQAALIIDGVFKSISDAVAGEEVS
IPGGKFKVQAKPRTGRNPTATITIAASKVGFPAKOLKDAVNG"
/3609..4838
/gene="CC2332"
/3609..4838
/gene="CC2332"
/note="Identified by Glimmer2; putative"
/codon_start-1
/transl_table-11
/product="conserved hypothetical protein"
/protein_id="AAK24303.1"
/db_xref="GI:13423857"
/translation="MLDMAYIDLRLKALISDAKYDASCASSAVORNSLNGKGVSS
TEGNGICHAIPDRCITSLKILTNFCITDCAICINRISNTFRAPFVOEVDTLL
NFKRNYIEGIFLSSGVIRSGDYEMEMVRAKSLRHDHFRGITHKLIPASPELA
AEGALDRAVSIINTELPRDSDIKALAEKDPDTRAMGRRLAIDHAEGRARRR
AFASOSTOILVGADEAKDGDILSRANLVGAYRLSRVYSAFSPIDPASRLPLVP
PLUREHLTYQADWIMRFYGFASQSEIIVAAGDLDLSVPKTSMAHRDRFPVDQIAD

```

```

gene
CDS
/note="Identified by match to TIGR protein family HMM
TIGR00758"
/gene
/codon_start-1
/transl_table-11
/product="phage SpoI DNA polymerase-related protein"
/protein_id="AAK24304.1"
/db_xref="GI:13423858"
/translation="MQVVRASEVDPAQWRRAARVIRACGAPESLWTVTERLND
EPYTDATTFVYPAAFEMAQOVLNRSIDRFLLRLIMRLREPRLEINPADQDMA
RADMAKAVSRAHKKKAFVRFVREDAKERTLFPDLFNGDRORRRRH
NMTILNDVCVANGERLWYSEGDADPADASEQALNMTYASIFNPARLRPMR
OEMPKRYVRLPEALIRPGIIEAODRAAMATPPRPSERYLKAORHARAPYNA
GAPITIDAVRAGVSVCRCDLREATQVPGDGAPSAPLMFVEQGDDEDLGSLY
GPAQGLEDRALAEVSVREQTVYTNVVKHKEHLNKRRLRHKPTQGEVSACRWMLDA
ERLVRPRTVYIMLGATPAAGVMGRPPVRETGRPLPLPDGVOGLVTFHPSYLLRLPD
AAARENAVRSFVVDLRLAGRIAGLAATSP"
/complement(6418..6494)
/gene="trna-his-1"
/product="trna-his-1"
/complement(6418..6494)
/gene="trna-his-1"
/complement(6418..6494)
/6557..7441
/gene="CC2334"
/6557..7441
/gene="CC2334"
/note="Identified by Glimmer2; putative"
/codon_start-1
/transl_table-11
/product="hypothetical protein"
/protein_id="AAK24305.1"
/db_xref="GI:13423859"
/translation="MNGKRPVETTPGSLTFVAYIOALLRDLERORGHATVUGRAVE
GLIHPEQAOADSRILAEAHVRRFGNHRARAFQSALRHGARDRATADHPLSTVRI
VPGHVFTRGLTGVTGRDORTSRVRRNRNRGRDRSRSLFPLFNGDRORRRRH
VRRSHRRARLNLNGRRRRRRTIGRRRRRRRRRRRROLHNRVYVLLHEHFEILL
ROISSOORVOTSHDGGDTFETLLATVAFRRERIRVAMLGHLLSHASPPGYHHY
RRAPRYRSRYPGCSPP"
/complement(7265..7799)
/gene="CC2335"
/complement(7265..7799)
/gene="CC2335"
/note="similar to SP:005606 GB:X70139 PID:49194;
identified by sequence similarity; putative"
/codon_start-1
/transl_table-11
/product="ExbD/TolR family protein"
/protein_id="AAK24306.1"
/db_xref="GI:13423860"
/translation="MAAKISGGGDRFNVNONSEINVTPEVDMLVLLIIFYAAFLA
SVSEVNLPPAVAKSSPPSPKPYVYSIKOSGLYGDNEISIDELGIDIKNNGRDP
SKRIRTRADEKRYKGFMEVANTLIDNGRTYVALYGRDSQ"
/complement(7805..8713)
/gene="CC2336"
/complement(7805..8713)
/gene="CC2336"
/note="Identified by match to protein family HMM"
/codon_start-1
/transl_table-11
/product="MocA/TolO/ExbB proton channel family protein"
/protein_id="AAK24307.1"
/db_xref="GI:13423861"
/translation="MLDIRKTFPILVGMALMASAPAFADAAAAPATAAOPAT
DAAAAPADAAAAPAEAPAEAPAAHSLTFPGMEQACVYVYVYVIGVLASY
FVSWLLTKLFEERALKKOSDKLEAVRGAHSIADRIKSSSEFBNPADVAAAAS
QVEVLISQOALVAGERHSDLSRAYAIVAAVOSLAKRLSGGVNVLASVSGSGPTIG
LFTVVGIGIMSFGIANTNTNLAVAVAGIAELALATGIGLFAAIPAVIFYNYNFORI
SAGTREGGVAAELMNAISRLDKGA"

```

gene		complement(9627 . 10454)	/gene="CC2337"
CDS		complement(9627 . 10454)	/gene="CC2337"
		/note="Identified by match to protein family HMM"	
		/codon_start=1	
		/transl_table=11	
		/product="oxidoreductase, short-chain dehydrogenase/reductase family"	
		/protein_id="AAK24308.1"	
		/db_xref="GI:13423862"	
		/translation="MSRSVKRSPRYGATLVGTASQSIGEPARLARSGYNLTLYA RRDRRLADAVOSKYCVQARVIAADDGRKEDYERYIAATAADYDGLIVAAAGFTGIS	
Query Match	11.1%	Score 66.4; DB 1;	Length 11684;
Best Local Similarity	54.0%;	Pred. No. 1.7e-05;	
Matches 136;	Conservative 0;	Mismatches 116;	Indels 0; Gaps 0;
Oy	227	tgcgccagaaatcccaactctgaccaaagtctcggcggaagctctgttaacgcttcasg	286
Db	3240	TGC GCCCGC GACCAAAAATCTGAAGAAGACCAGCCAGCATTCATCGACGCCGCTTCA	3299
Oy	287	atgttgtcgtcgaagctataagaatccggcggaagcgtgaagctaacggcgcttctccg	346
Db	3300	AGTCGATCTCCGACGCCGCCCGTGAAAGGGCGAGAGACTGTCAATCCC GGCTTCGCAACT	3359
Oy	347	ctgaagcgctcaagcgccgcgcgtctgcaccgcgcgaacccgcgcacttgtagcacattg	406
Db	3360	TCAAGTGTCAGAGCTTAAGCGCGGCCGACCGCGCCGACACCCGGCGACCGGGGACCATCG	3419
Oy	407	acattccggtctctctaaggcgtttgatctccgctgtgctccctgctgaagaagcgctga	466
Db	3420	AAATGCGCCGACGACAAGAGCTGGCTTCACGCCGCCGCAAGCACTGAAGACGCCGCTCA	3479
Oy	467	ccgaatgacctt 478	
Db	3480	ACGCGTAAAGCCT 3491	
RESULT 13			
AE004606			
LOCUS	AE004606	10851 bp	DNA
DEFINITION	Pseudomonas aeruginosa PA01, section 167 of 529 of the complete genome.		BCT 30-AUG-2000
ACCESSION	AE004606	AE004091	
VERSION	AE004606.1	GI:9947780	
KEYWORDS			
SOURCE			
ORGANISM			
	Pseudomonas aeruginosa.		
	Pseudomonas aeruginosa Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; Pseudomonas.		
REFERENCE			
AUTHORS	1 (bases 1 to 10851) Stover,C.K., Pham,X.-O.T., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S., Hutnagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Golty,L., Tolentino,E., Westbrock-Wadman,S., Yuan,Y., Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Larbig,K., Lim,R., Smith,K., Spencer,D., Wong,G.K., Wu,Z. and Paulsen,I.T.		
TITLE	Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen		
JOURNAL	Nature 406 (6799), 959-964 (2000)		
MEDLINE	20437337		
REFERENCE	2 (bases 1 to 10851) Stover,C.K., Pham,X.-O.T., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S.L., Hutnagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Golty,L., Tolentino,E., Westbrock-Wadman,S., Yuan,Y., Folger,K.R., Kas,A., Larbig,K., Lim,R.M., Smith,K.A., Spencer,D.H., Wong,G.K., Wu,Z., Paulsen,I.T., Reiker,J., Sater,M.H., Hancock,R.E.W., Iorv,s. and Olson,M.V.		
AUTHORS	Direct Submission		
TITLE	Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University Of Washington, Box 352145, Seattle, WA 98195, USA		
JOURNAL			

FEATURES

source	location/Qualifiers
gene	1..10851 /organism="Pseudomonas aeruginosa" /strain="PA01" /db_xref="taxon:287" complement(83..1369) /gene="PA1798" complement(83..1369) /gene="PA1798" /codon_start=1 /transl_table=11 /product="probable two-component sensor" /protein_id="AAG05187.1" /db_xref="GI:9947781" /translation="MRRLPRLVLLALAGFAAIFVVDHYIDAFYDSIVENYHDAVR GGAIVSYERKLAIPDQNGRORQLEDMKRPHTGLESLTDARQAKLTQEQALDLNKLAV KEPFEFISRIDGPDQLDKLPPEPSLPGLFTVLAVALIGVAVGIALVWVRHMD LETLRLAQRFGDGLDSRTIRIRSDIRTLAQHFQMDRISLISNQELTNVA SH ELNTPISRLSFELEQLNKQYDAVEVRDLIEDMDADIGELEEMVELTYARLEHGNG SHEEIVDASMLDVSADVADVALEAAAGVGEISACQVEQIRIRPRFAVITLARA IRRAHSRVEALILDDGDSQIRVNDGPGIPADARCKIFEPFRSLDSSRSTGCEFL GLAIVRAVMAHGCVAALETTPGCGASFLTLNRRP" complement(1370..2077) /gene="PA1799" complement(1370..2077) /gene="PA1799" /codon_start=1 /transl_table=11 /product="probable two-component response regulator" /protein_id="AAG05188.1" /db_xref="GI:9947782" /translation="MDCPILSKVLVLEDPOKTLARLIASFLOGHFEVRYHGRDAAPA ATLDEPQVYVLDLMPGONGLOVCEIRRVANPLITLTAQEDDLHILGESSGD VYKPEPVLTLARLALMRRLRPASPESITFGKLTNRREARLBSGLIELTM EHELMLLPSQAGEILSRBEILNQITNGITGFDGILNSYVCISLKRKLKDNPREPVR KTVWGGYLPNPLGWL" 2309..3619 /gene="tig" /note="PA1800" 2309..3619 /gene="tig" /codon_start=1 /transl_table=11 /product="trigger factor" /protein_id="AAG05189.1" /db_xref="GI:9947783" /translation="MQVSYESTSALREMTYGPAPRIETEVNKRILQQTARAKITGF RGKYPMSVIRORYTESARQEMAGDILQETTFTEAVYEQKLNPGSGSVESKSEKND LEYIATFEPEFIVSGLEDIKVERLQAVASDVADNMDLNKLQREVEVERAQN DQNLIDFVKGIDGFAFAGSAGAKGLLVAGSGMTLGFEGVLGAKAGERVYLNTPN EDYQNLIDLANKAAEFTVYVNAEPKLPENEFEPALFGVKEGTGLGFAEYQKNMER EYRQAKRSKYNQNMVEGLQANPIEYKRALIGNEVRLRVQAVQOEGGNIKPOLPAE LEEBQAKRSEVGLIYAEVYKQHELKADGGRVREMTLEMASAVQEPQVYVNTFKNEP QINERSVLIEQGVVDVYIQKATVYDKQVSYEAYKAPAAPAA" 3713..4354 /gene="clpX" /note="PA1801" 3713..4354 /gene="clpX" /codon_start=1 /transl_table=11 /product="ATP-dependent Clp protease proteolytic subunit" /protein_id="AAG05190.1" /db_xref="GI:9947784" /translation="MSRNSFIPHPVDIQAGGLVPVVRQSGRGERAYDIYSLLKER IIFLGOVEDYKANTVVAQLFLLEAENPEKDIHLYINSGSGSVTGMSTYDPMQFKP NVSTTCIOACSMGALLLAGAAGKRYCCLPHRMATIHOPLGFGQASDIETLHAKETIL FIKELNLQILAHHTQPLDIYARDYDRFRMSGDEAVKIGLIDKWTQDRLAV" 4459..5739 /gene="clpX" /note="PA1802" 4459..5739 /gene="clpX"

[illegible]

BASE COUNT	2225 a	3228 c	3439 g	1959 t	GDYVIRLNGVSEEEAASIDSEKAMRYRFLASRSGQADFAFRQLQKAEVEXY"
ORIGIN					
Query Match	10.8%	Score 64.6;	DB 1;	Length 10851;	
Best Local Similarity	52.0%;	Pred. No. 4.6e-05;			
Matches 145;	Conservative 0;	Mismatches 134;	Indels 0;	Gaps 0;	
Oy	202	aacaagcttgacatctcggtttcgaagatcgccagaagtcacacattaccagaagctcaagcc	261		
Db	8406	AACAAGTGGGAACGTATGATGATGACAAATTCGCCGATCGTGTGATATCCGAAACCTGTGGC	8465		
Oy	262	gaagctgtctgtaaacgccttcacagatgtgtctgcgcgagctatgaagtcgcgcgaagc	321		
Db	8466	GGTGCCGCACTGAGACGCAAGTATGATGACAGTCGTCACATGGCGCCCTTAAGGCTGTACTCTC	8525		
Oy	322	ctgaagctaccagcgcctgtttctcgcgttagcgcgtcaagcgcgcgcgttcgcacgcgc	381		
Db	8526	GTGCGTGTGGTGTGGTTCGGACCTTCCTGTCACAAAGAGCGCGCTGCTGCACCGGTGCG	8585		
Oy	382	aaccgcgcactgagcgcagcagattgaacatccgcgttcctacgcggttcgtatctccgct	441		
Db	8586	AACCGCAACTGCGACACCCGATCAAGATGCGCGCTGCCAAGATCCCGGCTTCAAGCC	8645		
Oy	442	ggctccctgctgaagaagcgcgctacgcagtagtacctct	480		
Db	8646	GGTAAAGCCCTGAAGGATGCTGTCAACATCAAGCGCTTTT	8684		
RESULT 14					
AE004115					
LOCUS	12116 bp	DNA	BCF	31-JUN-2000	
DEFINITION	Vibrio cholerae chromosome I, section 23 of 251 of the complete				
ACCESSION	AE004115				
VERSION	AE004115.1				
KEYWORDS	GI:9654674				
SOURCE					
ORGANISM	Vibrio cholerae.				
REFERENCE	Vibrio cholerae.				
AUTHORS	1 (bases 1 to 12116)				
	Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwin, M.L.,				
	Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.A.,				
	Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,				
	Ernst, M.D., Vamathevan, J., Bess, S., Qin, H., Dragoi, I.,				
	Sellers, P., McDonald, L., Utterback, T., Fleischmann, R.D.,				
	Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,				
	Mekalanos, J.J., Venter, J.C. and Fraser, C.M.				
	DNA sequence of both chromosomes of the cholera pathogen Vibrio				
	cholerae				
	Nature 406 (6795), 477-483 (2000)				
JOURNAL	2 (bases 1 to 12116)				
MEDLINE	20406833				
REFERENCE	1 (bases 1 to 12116)				
AUTHORS	Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwin, M.L.,				
	Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.A.,				
	Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,				
	Ernst, M.D., Vamathevan, J., Bess, S., Qin, H., Dragoi, I.,				
	Sellers, P., McDonald, L., Utterback, T., Fleischmann, R.D.,				
	Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,				
	Mekalanos, J.J., Venter, J.C. and Fraser, C.M.				
	Direct Submission				
TITLE	Submitted (14-JUN-2000) The Institute for Genomic Research, 9712				
JOURNAL	Medical Center Dr, Rockville, MD 20850, USA				
FEATURES	Location/Qualifiers				
SOURCE	1..12116				
	/organism="Vibrio cholerae"				
	/strain="N1961"				
	/serotype="O1"				
	/db_xref="taxon:666"				
	/chromosome="I"				
	/note="Biotype: El Tor"				
	complement(167..1400)				
	/gene="VC0265"				
gene					

CDS complement(162. .1400)
 /gene="VC0265"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF93440.1"
 /db_xref="GI:9654675"
 /translation="MLETITLMLKFDIEIRVIQALAEALGLIAISIAIHILVLRGV
 LLMKRTTHHQSWMNALFEKHELSFVALLIQGVAVIIVQDMLSPMTFADVLT
 TSLIVVEGMATYSLVNLLEISRGVGNLPRGIIOSIKLIVFIITALLFFSII
 IGRSPVLLSGIGAMTAVMLVFEPDIPGLVAGVQSLSNKMLSVGDMLEMYKGDGS
 VIDLSITVKKQNMKDTITTPSYALISDSFRNMKGQSGGRRIKSLIDATSVHF
 LTSEKAKLAQQLLEPYLVEKEQEISSYNQKMDLACRINGRRLTNGISRAVLR
 YLRHPIHDMQTLMLEROLAPTHDGISLEVCSTVWVEYERIQSDIFDIYAVLP
 EFDLRVQAPGNFRLRG"
 complement(1455. .2027)
 /gene="VC0266"
 complement(1455. .2027)
 /gene="VC0266"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF93441.1"
 /db_xref="GI:9654676"
 /translation="MLESLLTEIRQCTVCEPHLPGLGPNVIRAHPAKILIIQGNPFT
 KVIQTSIPMDNASERLQWGLDREAFYCEENAIIMPGLCYPGKSGSDLPKREC
 APIHMAVLVRLPRLQTLILIGYAOHYYLSDKSTLTLETYQGMRAAPSLPLPHPS
 PRNTLMKKNPMEFQDIPYIRORVQVLT"
 complement(2027. .2950)
 /gene="VC0267"
 complement(2027. .2950)
 /gene="VC0267"
 /note="similar to GB:z11929 SP:p43169 PID:46871;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="D-isomer specific 2-hydroxyacid dehydrogenase
 family protein"
 /protein_id="AAF93442.1"
 /db_xref="GI:9654677"
 /translation="MNNFHOLYITDEDEHYREQILEROLEGALTEDRAKATILLAA
 PPAARALDEPRLQMLHSAYAGVDTMDPRLRNLYLTNVKGIFFGLIAYVYAI
 QYQDFRLYQTOAERLQMPRPSSIANQTLVIGTGSIGSHLAHVAKQFGIVAGVN
 RTGIPAKEGHGDATYHISELPALMARADLLVNTLPMTPATGSLIHOENLRCHQALLP
 NVGRGTVROGLPDLIAAGHIRAFDVPYIKKEPLADDPMDNPALITTPHIAVSF
 PRQVVDIFRNYORWCDNLPDRNIDDEKGT"
 2943. .3611
 /gene="VC0268"
 2943. .3611
 /gene="VC0268"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF93443.1"
 /db_xref="GI:9654678"
 /translation="MTIGASFPPGYVHFRCLFPQREVTNLONPQLRLKEKEPWO
 QITFACLCRRMTPTMTAFCEHDFQFAEARTYRDLDSVWELATVKNNAVNEHOLEKL
 EELIIPSDADLAVYPAIDACBGLATLGLDRDLAESEMIKVSQISQVQALVE
 AQCEGKITNDNORENEAVCAEMDVMAIFRPLREAVSRDIDLKDLKRLKEEGSVNI
 GVEL"
 3974. .5173
 /gene="VC0269"
 3974. .5173
 /gene="VC0269"
 /note="similar to SP:p25081 GB:x57117 PID:48820;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="mannose-6-phosphate isomerase"
 /protein_id="AAF93444.1"
 /db_xref="GI:9654679"
 /translation="MNTAIPADPFFIMHANYIQSYAMQSDISINOLFDPINPKGEPOAV
 NMGTHPNCGSYVESEGIKIPDLIOKRTDPSIQTVQFGLPFLYFKLIILANOALS
 IOVHPKAEFGFAEENAGNGLIINATNNYKDPNKRPELYVALTINQANGRPLSD
 IISPSALAIIDELASILSKRONQTCGHEFTKLSLCKGKRRVLEQLSYASH
 AAOPIPALIEELATQTPNDIGLAPMLLVLTQPEANMLDARTPAVLYKGLATJM
 ANSDVNLRAGLTPKTHIDVELAKCTFEKSEKATLLTPQCDMSLFPVADFKFA
 IPRPOQANITVSSAELLPIQCDATVSPCGVTLVHHGQSVFIPAYTEKYCLSAAG
 RVARAVN"
 5139. .6119
 /gene="VC0270"
 5139. .6119
 /gene="VC0270"
 /note="similar to SP:p23917 GB:x76979 PID:145329
 PID:58151 GB:U00096; identified by sequence similarity;
 putative"
 /codon_start=1
 /transl_table=11
 /product="ROK family protein"
 /protein_id="AAF93445.1"
 /db_xref="GI:9654680"
 /translation="MQQVDSREHTINNGSSMADHDEHPTEQKFFYMLIGLDTGTEIE
 ICYLDKQSMLYRQRTAPDNTSQPDVCSLVDPEQATQPDVDSIGILPGVSPVT
 GLIKNANCTEVLNGODLSQLYRLGREVYLDANCLPSEALIDAGKESMVVEGAIL
 GTGCGSIVYVNOVLYVPAVNAICGEMHNLPGLHLEQDGAARCYCGRQNCIERFTSG
 SGFQDSYQALTEGCITASEIMKRYKQDEALHCYQLIDHMAKSPAGLVNVDPII
 VLGGLSNIDELVRLPLPTARTAHVPSDSQVHFAKAVPDDSSIRGANML"
 complement(6467. .7525)
 /gene="VC0271"
 complement(6467. .7525)
 /gene="VC0271"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF93446.1"
 /db_xref="GI:9654681"
 /translation="MELAIYISVAIGVSFICSVLEAVLSTIPSYLQLRQGHRA
 NMLAGLKIDIPRLSITITNFIARTIGATAGAAVYFGSQZLFENAVLITGLV
 LSEIYPTKIGATYRELPQASLVLRMHWALTPFWSEQLTKRLARVEAKLDE
 ISAMALANNEEFAGESEKMLNNLAIQNVQVQVMPVLPFRVSAQLTIDELFQ
 HNDPFSRPLIYSEKDNVGVFVHLEFKEQONGNLLLDVMPHIVAVNTLSLP
 KAFDMQMRQLQSVYVDEYGSVQGLTLEDIFEHILGSEIIDEADRTDMQOLATER
 WEHMRORMISRDEVE"
 7534. .7698
 /gene="VC0272"
 7534. .7698
 /gene="VC0272"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAF93447.1"
 /db_xref="GI:9654682"
 /translation="MYLKTFFELFSSHNLKGRKAQGEYEKSLRKPKYLVSVTK
 SIEMGFPVD"
 7897. .8169
 /gene="VC0273"
 7897. .8169
 /gene="VC0273"
 /note="similar to SP:p28080 PID:141878; identified by
 sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="DNA-binding protein HU"
 /protein_id="AAF93448.1"
 /db_xref="GI:9654683"
 /translation="MNTKQLIDFIEKADLTGVQAKALEATIGAVEGALKRGDQVQL
 IGFGRYKNHNSARGRNPKGEELIKIAANVPAVACKALKDAIK"
 8170. .8904
 /gene="VC0274"
 8170. .8904
 /gene="VC0274"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="mannose-6-phosphate isomerase"
 /protein_id="AAF93449.1"
 /db_xref="GI:9654684"
 /translation="MNTAIPADPFFIMHANYIQSYAMQSDISINOLFDPINPKGEPOAV
 NMGTHPNCGSYVESEGIKIPDLIOKRTDPSIQTVQFGLPFLYFKLIILANOALS
 IOVHPKAEFGFAEENAGNGLIINATNNYKDPNKRPELYVALTINQANGRPLSD
 IISPSALAIIDELASILSKRONQTCGHEFTKLSLCKGKRRVLEQLSYASH
 AAOPIPALIEELATQTPNDIGLAPMLLVLTQPEANMLDARTPAVLYKGLATJM
 ANSDVNLRAGLTPKTHIDVELAKCTFEKSEKATLLTPQCDMSLFPVADFKFA
 IPRPOQANITVSSAELLPIQCDATVSPCGVTLVHHGQSVFIPAYTEKYCLSAAG
 RVARAVN"
 8170. .8904
 /gene="VC0274"
 8170. .8904
 /gene="VC0274"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="mannose-6-phosphate isomerase"
 /protein_id="AAF93450.1"
 /db_xref="GI:9654685"
 /translation="MNTAIPADPFFIMHANYIQSYAMQSDISINOLFDPINPKGEPOAV
 NMGTHPNCGSYVESEGIKIPDLIOKRTDPSIQTVQFGLPFLYFKLIILANOALS
 IOVHPKAEFGFAEENAGNGLIINATNNYKDPNKRPELYVALTINQANGRPLSD
 IISPSALAIIDELASILSKRONQTCGHEFTKLSLCKGKRRVLEQLSYASH
 AAOPIPALIEELATQTPNDIGLAPMLLVLTQPEANMLDARTPAVLYKGLATJM
 ANSDVNLRAGLTPKTHIDVELAKCTFEKSEKATLLTPQCDMSLFPVADFKFA
 IPRPOQANITVSSAELLPIQCDATVSPCGVTLVHHGQSVFIPAYTEKYCLSAAG
 RVARAVN"
 8170. .8904
 /gene="VC0274"
 8170. .8904
 /gene="VC0274"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="mannose-6-phosphate isomerase"
 /protein_id="AAF93451.1"
 /db_xref="GI:9654686"
 /translation="MNTAIPADPFFIMHANYIQSYAMQSDISINOLFDPINPKGEPOAV
 NMGTHPNCGSYVESEGIKIPDLIOKRTDPSIQTVQFGLPFLYFKLIILANOALS
 IOVHPKAEFGFAEENAGNGLIINATNNYKDPNKRPELYVALTINQANGRPLSD
 IISPSALAIIDELASILSKRONQTCGHEFTKLSLCKGKRRVLEQLSYASH
 AAOPIPALIEELATQTPNDIGLAPMLLVLTQPEANMLDARTPAVLYKGLATJM
 ANSDVNLRAGLTPKTHIDVELAKCTFEKSEKATLLTPQCDMSLFPVADFKFA
 IPRPOQANITVSSAELLPIQCDATVSPCGVTLVHHGQSVFIPAYTEKYCLSAAG
 RVARAVN"
 8170. .8904
 /gene="

```

/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF93449.1"
/db_xref="GI:9654684"
/translation="MTLPRLNINSAFLKWKQSLFLSLGLAGCSSTPNPNLEQD
NLPFGKGTIGDNTSFYWTTSFQKPSASLYTSDGYGYSQSYTHMEGQVREIRRE
EHLGKGLKIPVRAHRIKFKSGEAVYQYRLRGVGLVPMNEQLAHVLDAAKVAETKE
ODKGLKLIQGYNMGKFLKSGEAVYQYRLRGVGLVPMNEQLAHVLDAAKVAETKE
GVHIDELLLDLDGHCVCVEPELID"
/codon_start=1(9004..10293)
/gene="VC0275"

```

Query Match	10.7%;	Score 64.4;	DB 1;	Length 12116;
Best Local Similarity	51.8%;	Pred. No. 5.1e-05;		
Matches 146;	Conservative 0;	Mismatches 136;	Indels 0;	Gaps 0;

QY	202	aacaaatctgacgttcgttttcgaagatcgcgccgaagtcacattacccaagctcaagcc	261
Db	7900	AACAAGACCCTATTCATGACTTTTTCGCAGAGAAAGGAGACCTGACATCAATGCAAGC	7959
QY	262	gaagctcgtcttaacgcctccagaatgtytctgcgaagctatgaagtcggcgaaagcc	321
Db	7960	AAAGACGCTCTTGAAGCAGCTCTGGGGTCCGGTTGTAAGGTCCTGAAGAAGCAGCACCA	8019
QY	322	ctgaagctccacgcgcctctctcgcctgagcgcgcgaagcgcgcgcgcgcgcgcgcgc	381
Db	8020	GTACACATCGATTGGTTTGTGTCATCTCAATCAAGTTAATCATTCGTTCTGCTGTGACTGGCCT	8079
QY	382	aaccgcgcactcggcgaagcagatgacattccgcgcttcctacagcggttcgattctccgct	441
Db	8080	AATCCAAAGACCGCGGAGAGATCAAAATTTGCGGACAGTCAAGCTTCCTCATTTCTGACGA	8139
QY	442	ggtccctcgtctgaagaagcgcgcctcccgagctctcgtc	483
Db	8140	GGTAAAGCGCTGAAGAAGCGCATCAAGTATATGCTTTTATCT	8181

BASE COUNT

ORIGIN

758 a 971 c 1037 g 614 t

gene

CDS

transl_table=11

/product="lon protease"

/protein_id="AA05300.1"

/db_xref="GI: 6175841"

/translation="MTNKISIPATESATYEVLPFLDLYVEFPMHIVPLFVGRKSTRALAE
EWGTDQIMLVYQINATDDPEPSAIYKGTIANVQLKLPDGYTKVIVEGRSRAE
IEERTPDDEVEAMAHALPEPDEPVEILETSSVSEFESYVFLANKIIPREVGVAS
QIDYRSLADTVASHLSIKIYVKOEIMETTSYKMRKALGFMEGELISVLOVERKAS
RIVROMEKTPOREYLIYNOMKAIORKEIMSDSDGDEMELBERISKTSEARERKAA
ELKIKRQMSMAAPATVPRYLDIMLGLDIPMGKSKITDQINHAKEYIDPHGICIDYK
ERIVETIYAVARBSKIKGPIYPLICFVGPVGVSTSAKSIANATGRETIRALAGVDRDA
EIRGHRRYTGSMFGKVVQSMKRAKSNPYFLDEIDDKQDDEGDSALLEVYDA
QNSTFMHYLEVELDLSNMEITTYANTLIINPPLMDMEVFRLAGYEEERLLETAKH
LLEKRAIDHALQNEFSEVTDGALTALIONTREGVSNPRELAKRLAVETLIRK
TKREYSAENINIDYLVPRRGEAEEDDQGVVTCGLATMEVGGELLITIGVAMPFGK
RMVGTGLRDVYKESIIASAASYRSLIDGIEPDEDKDIDHNVHEGATPKNGPSA
GVMAATVYISVMTGIPISKIDVAMTGETTLINGVLIIGGLEKILLALRGGIIKTVLIDP
ENAKDIDIPDYNVRKNSFEIIPVSRMGEYVHALIRLEPIEMDPASPALPVSQSD
ENAGISINH"

2992. .3264

/gene="hupB"

2992. .3264

/gene="hupB"

/note="hupB"

/codon_start=1

/transl_table=11

/product="histone-like protein"

/protein_id="AA05301.1"

/db_xref="GI: 6175842"

/translation="MNKNEIYAAVDKAGLSKADASSAVDAVFETIOGELKNGGDIYLT
VGCNFSVSRKASKGNSPTGAEVYDIPARNVPRFTMGKGLKDAVN"

Query Match	Similarity	10.68	Score 63.8	DB 1	Length 3380
Best Local Match	Similarity	52.48	Pred. No. 8e-05		
Matches 140	Conservative	0	Mismatches 127	Indels	Gaps
QY 202	aacaagctgacccctcgtttcttgcgaagatcgcgcccaagaagtcacaactgcattaccagaagctcagcgc	261			
Db 2995	AACCAAAATGAGCTCGTGGCTGCCGTACCCGACCAAGGCCGGGACTCTGCAGAGGCAGATCGC	3054			
QY 262	gaagcgtgctgtttaaagcgccttcacgaatgtgttcgtcgaagctatgaagtcagcgcaagcgc	321			
Db 3055	TCTTCTGCCGTTGATGACGAGTTTTTGAGACCAFTTCAGGGCGCAACTGAAAGACGGCGCGCAC	3114			
QY 322	ctgaagctacacgcgcctgtttctccgctctgaagcgcgcgtcaaaagcgcgcgcgcgcgcgcgcgc	381			
Db 3115	ATTCGGCCCTGTTGGCGCTTTCGGCAATTTCCCTATTCGGCCGCCGCGAGGCTTACGAAGGGCCGC	3174			
QY 382	aaccgcgcactgcgcgcgcagattgaaattccgcgttctctatcgcgcgttcctatctccgcct	441			
Db 3175	AACCGCTGACGGGGTCCGGAAGTTGACATTCGCCGACCAAGTCCGAGATTCACGGGCC	3234			
QY 442	ggctcccgctctgaagaagcgcgtcacc	468			
Db 3235	GGCAGGCGCTTCAGAGACCGCGTCAAC	3261			

Search completed: December 20, 2001, 19:47:34
Job time: 6535 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2001, 18:54:44 ; Search time 167.05 seconds
(without alignments)
3079.288 Million cell updates/sec

Title: US-09-816-391a-1

Perfect score: 600
Sequence: 1 gctggcgccgagcgccatga.....tattccgactagtcagcg 600

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq.1101.*
1: /SID8/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID8/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID8/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID8/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID8/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID8/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID8/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID8/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID8/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID8/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID8/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID8/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID8/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID8/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID8/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID8/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID8/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID8/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID8/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID8/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID8/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID8/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	47.2	7.9	349980	21	AAAF21609
C 2	47.2	7.9	349980	21	AAAF21610
C 3	47.2	7.9	1437668	21	AAA81490
C 4	46.8	7.8	8093	21	AAA81512
C 5	45	7.5	62909	22	AAAF28545
C 6	43.6	7.3	16075	20	AAAF99811
C 7	43.2	7.2	16079	8	AAAF70753
C 8	40.6	6.8	273	21	AAAF79661
C 9	39.6	6.6	1026	21	AAA88001
C 10	39.6	6.6	1575	15	AAAF0417
C 11	39.6	6.6	1578	15	AAAF0418

12	39.6	6.6	1638	17	AAAF9320
13	39.6	6.6	1638	22	AAAF27164
C 14	39	6.5	1326	22	AAAF4356
C 15	39	6.5	1326	22	AAAF48225
C 16	39	6.5	1326	22	AAAF49591
17	37.8	6.3	276	22	AAAF90714
18	37.8	6.3	276	22	AAAF90729
19	37.6	6.3	858	20	AAAF11393
20	37.4	6.2	969	20	AAAF98059
C 21	37.4	6.2	1017	20	AAAF98058
22	37.4	6.2	42235	20	AAAF98035
23	37	6.2	2943	20	AAAF16480
24	36.6	6.1	77536	21	AAAF14651
25	36.2	6.0	300	22	AAAF90842
26	36.2	6.0	375	20	AAAF98061
27	35.8	6.0	2025	20	AAAF98060
28	35.4	5.9	1822	11	AAAF06095
29	35.2	5.9	683	21	AAAF13147
30	35.2	5.9	47981	22	AAAF30757
31	35	5.8	1950	19	AAAF23479
32	35	5.8	1950	21	AAAF13897
33	35	5.8	3246	18	AAAF93610
C 34	35	5.8	17612	19	AAAF23494
C 35	35	5.8	17612	21	AAAF13905
C 36	35	5.8	17612	22	AAAF30870
C 37	34.8	5.8	2721	12	AAAF01212
C 38	34.8	5.8	2721	12	AAAF01543
C 39	34.8	5.8	8438	15	AAAF073500
40	34.6	5.8	624	17	AAAF06312
C 41	34.6	5.8	2001	22	AAAF60958
C 42	34.6	5.8	10732	21	AAAF10594
C 43	34.4	5.7	1524	12	AAAF01499
44	34.4	5.7	1525	20	AAAF91207
45	34.2	5.7	471	21	AAAF15020

ALIGNMENTS

RESULT 1	AAAF21609/C	AAAF21609 standard; DNA; 349980 BP.
ID	AAAF21609	
AC	AAAF21609;	
XX		
DT	13-MAR-2001	(first entry)
XX		
DE	Neisseria meningitidis B nucleotide sequence SEQ ID NO:110.	
XX		
KW	Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine; diagnosis; antigen; detection; infection; gene therapy; antibacterial; ds.	
XX		
OS	Neisseria meningitidis.	
XX		
PN	WO200066791-A1.	
XX		
PD	09-NOV-2000.	
XX		
PF	08-MAR-2000; 2000WO-US05928.	
XX		
PR	30-APR-1999; 99US-0132068.	
PR	08-OCT-1999; 99WO-US23573.	
PR	28-FEB-2000; 2000GB-0004655.	
XX		
PA	(CHIR) CHIRON CORP.	
PA	(GENO-) INST GENOMIC RES.	
PI	Piazza M, Hickey E, Peterson J, Tettelin H, Venter JC, Maignant V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappelli R, Frazer CM, Grandi G;	
XX		
XX	WPI; 2000-647603/62.	

Cholesterol oxidase
Streptomyces preg
Human GAT4 nucle
Heart muscle cell
Human GAT4 codin
CPE 14 codin sequ
CPE 13 codin sequ
M. vaccae antigen
Nucleotide sequenc
Nucleotide sequenc
SAB virus gB glyco
Nucleotide sequenc
2CPE 33 codin sequ
Nucleotide sequenc
Nucleotide sequenc
Sequence encoding
Aspergillus oryzae
Microsporidia meg
Pseudomonas xcp s
Pseudomonas alcali
Mycobacterium tube
Pseudomonas alcali
Pseudomonas alcali
BamHI J-I fragment
BamHI J-I fragment
DNA encoding Pseud
Thaumatin II synth
P. putida KT2440-a
Gene encoding a su
Beta-ketothiolase
Beta-ketothiolase
Tritichoderm reesel

XX Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections -
XX
PS Claim 7: Appendix A: 692pp; English.

XX The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.

XX Sequence 349980 BP; 81351 A; 86755 C; 95584 G; 86230 T; 0 other;

SO Query Match 7.9%; Score 47.2; DB 21; Length 349980;
Best Local Similarity 48.8%; Pred. No. 0.0052;
Matches 127; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 202 aacagctcgactcggttcgaagatcgccagagatccaactgcaagctcagcgc 261
DB 335204 AACCACTCGAATTCGATCGAAGGATGTCGCAAGAGCCGACATTTCCAAAGCCGCCGCA 335145
QY 262 gaggcctgcttaacgacctccagagatgctgctgagagctatgaagtcgcaagc 321
DB 335144 CAAAAAGCTTTGGATGCCACTACCAATGCAAGTACCAACGCCCTGAAACAAGCGCACAC 335085
QY 322 ctgaagctacacgacctgtcttcgcgtgagcgctcaagcgccgctcgacggcgccg 381
DB 335084 GTTACTTGTGCGTTCGCTTCTACGTGCGGCAACGTCGGAACCGCAAGGCCGCGC 335025
QY 382 aacccgacgacgtgagcagatgacatccggtctcctcagcggtctgatactcgcgt 441
DB 335024 AACCCCAAAACCGCGACGCTCTGACCAATTCGCGCCGCAAAAGCCCTAAATTCGCGGCC 334965
QY 442 ggctccctgctgaagaagc 461
DB 334964 GGCAAAGCTTTGAAAGACGC 334945

RESULT 2
AAF21610/c
ID AAF21610 standard; DNA; 349980 BP.
XX AAF21610;
XX 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:111.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
ds.
XX
OS Neisseria meningitidis.

PN WO200066791-A1.
XX
PD 09-NOV-2000.
XX
PF 08-MAR-2000; 2000WO-US05928.
XX
PR 30-APR-1999; 99US-0132068.
PR 08-OCT-1999; 99WO-US23573.
PR 28-FEB-2000; 2000GB-0004695.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Pizza M, Hickey E, Peterson J, Tetteijn H, Venter JC, Masignani V,
PI Galositi C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappelli R;
PI Frazer CM, Grandi G;
DR WPI: 2000-647603/62.
XX
XX Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections -
XX
PS Claim 7: Appendix A: 692pp; English.

XX The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.

SO Query Match 7.9%; Score 47.2; DB 21; Length 349980;
Best Local Similarity 48.8%; Pred. No. 0.0052;
Matches 127; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 202 aacagctcgactcggttcgaagatcgccagagatccaactgcaagctcagcgc 261
DB 335204 AACCACTCGAATTCGATCGAAGGATGTCGCAAGAGCCGACATTTCCAAAGCCGCCGCA 335145
QY 262 gaggcctgcttaacgacctccagagatgctgctgagagctatgaagtcgcaagc 321
DB 335144 CAAAAAGCTTTGGATGCCACTACCAATGCAAGTACCAACGCCCTGAAACAAGCGCACAC 335085
QY 322 ctgaagctacacgacctgtcttcgcgtgagcgctcaagcgccgctcgacggcgccg 381
DB 335084 GTTACTTGTGCGTTCGCTTCTACGTGCGGCAACGTCGGAACCGCAAGGCCGCGC 335025
QY 382 aacccgacgacgtgagcagatgacatccggtctcctcagcggtctgatactcgcgt 441
DB 335024 AACCCCAAAACCGCGACGCTCTGACCAATTCGCGGCCGCAAAAGCCCTAAATTCGCGGCC 334965
QY 442 ggctccctgctgaagaagc 461
DB 334964 GGCAAAGCTTTGAAAGACGC 334945

RESULT 3
 AAA81490/c
 ID AAA81490 standard; DNA; 1437668 BP.
 AC AAA81490;
 XX
 XX
 DT 04-DEC-2000 (first entry)
 DE N. meningitidis B full length genome DNA sequence SEQ ID NO:1068.
 XX
 XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
 KM antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KM Meningococcus B; MenB; ds.
 XX
 XX Neisseria meningitidis.
 OS
 XX WO20002430-A2.
 PN
 XX 20-APR-2000.
 PD
 XX 08-OCT-1999; 99WO-US23573.
 XX
 XX 09-OCT-1998; 98US-0103794.
 PR 30-APR-1999; 99US-0132068.
 XX
 XX (CHIR) CHIRON CORP.
 XX
 XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
 PI Mesigian V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappuoli R, Pizza M;
 PI
 DR WPI: 2000-318079/27.
 XX
 XX Isolated nucleotide sequences of Neisseria meningitidis which can be
 PT used in the diagnosis and treatment of N. meningitidis infection and
 PT other Neisserial infections, for example, N.gonorrhoea -
 XX
 PS Claim 7; Page 866-1272; 1760pp; English.
 XX
 XX The present invention describes methods of obtaining immunogenic
 CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
 CC represent specifically claimed Neisseria meningitidis genomic DNA
 CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
 CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
 CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
 CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
 CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
 CC sequences, which are all used in the exemplification of the present
 CC invention. The nucleic acid sequences, protein sequences, and antibodies
 CC against them, can be used in the manufacture of a composition. The
 CC composition can be used as a medicament (or in the manufacture of a
 CC medicament) for treating, preventing or diagnosing infection due to
 CC Neisserial bacteria. For example, some of the identified proteins could
 CC be components of vaccines against Meningococcus B; against all serotypes;
 CC and/or against all pathogenic Neisseriae. Identification of sequences
 CC from the bacterium will also facilitate production of biological probes,
 CC particularly organism-specific probes. Attempts to make efficacious
 CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
 CC Multivalent vaccines have also been tried but none have successfully
 CC overcome antigenic variability. The provision of further, complete
 CC sequences may provide an opportunity to identify secreted or surface
 CC exposed proteins that may be presumed targets for the immune system and
 CC which are not antigenically variable or at least more conserved than
 CC other more variable regions.
 CC
 CC Sequence 1437668 BP; 344338 A; 353206 C; 385074 G; 355045 T; 5 other;

Query Match 7.9%; Score 47.2; DB 21; Length 1437668;
 Best Local Similarity 48.8%; Pred. No. 0.0097;
 Matches 127; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

OY 202 aacaagctgcactgctgttcgaagatgcgccagaatccaaactgaccgaagctcagcc 261
 |||||
 Db 1205593 AACAAAGCTTGAAATGATGCAAGACGATTGCTCAGAACCCGACATTTCAAGGCCGCCA 1205534
 OY 262 gaggcctgcttaacgcttcaccagatgtgttcgtcgaagctatgaagctccgcaagc 321
 |||||
 Db 1205533 CAANAAGCTTTGGATGCGCATACCAATGCAATGTAACCAACCCCTGGAACAAGCGACAC 1205474
 OY 322 ctgaagctcaccgctgttctccgctgaagcgcgtcaagcgccgctgcagccgcccgc 381
 |||||
 Db 1205473 GTTACCTTGGTCGGTTCCGCTACTTCTACGTGGCGGCAACGTCCGAAACCAAGCCGC 1205414
 OY 382 aaccgcgactggcgaagcagattgacattccgcttcctcctcgcgcttcgatacgcgt 441
 |||||
 Db 1205413 AACCCCAAAACCGCGAGCCTTGACCATTTGCCGCCCAAAAGCCTTAATTCGCGGCC 1205354
 OY 442 ggtccctcgtcgaagaagc 461
 |||||
 Db 1205353 GGCAAAGCTTTGAAGACGC 1205334
 RESULT 4
 AAA81512
 ID AAA81512 standard; DNA; 8093 BP.
 XX
 XX AAA81512;
 AC
 XX
 DT 04-DEC-2000 (first entry)
 DE N. meningitidis partial DNA sequence gum_59 SEQ ID NO:59.
 XX
 XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
 KM antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KM Meningococcus B; MenB; ds.
 XX
 XX Neisseria meningitidis.
 OS
 XX WO20002430-A2.
 PN
 XX 20-APR-2000.
 PD
 XX 08-OCT-1999; 99WO-US23573.
 XX
 XX 09-OCT-1998; 98US-0103794.
 PR 30-APR-1999; 99US-0132068.
 XX
 XX (CHIR) CHIRON CORP.
 XX
 XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
 PI Mesigian V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappuoli R, Pizza M;
 PI
 DR WPI: 2000-318079/27.
 XX
 XX Isolated nucleotide sequences of Neisseria meningitidis which can be
 PT used in the diagnosis and treatment of N. meningitidis infection and
 PT other Neisserial infections, for example, N.gonorrhoea -
 XX
 PS Claim 7; Page 1372-1375; 1760pp; English.
 XX
 XX The present invention describes methods of obtaining immunogenic
 CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
 CC represent specifically claimed Neisseria meningitidis genomic DNA
 CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
 CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
 CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
 CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
 CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
 CC sequences, which are all used in the exemplification of the present
 CC invention. The nucleic acid sequences, protein sequences, and antibodies
 CC against them, can be used in the manufacture of a composition. The
 CC composition can be used as a medicament (or in the manufacture of a

PT diagnosis of endocrine diseases
 XX
 PS Claim 7; Page 9-11; 15pp; Japanese.
 XX
 CC Cholesterol oxidase is used for measurement of cholesterol which is
 CC an index for the diagnosis of endocrine diseases and abnormal
 CC metabolism. The enzyme is obtained by adding amino acid residues to
 CC the N-terminus of cholesterol oxidase derived from Streptomyces sp.
 CC SA-COO, which shows the same activity as that of the naturally
 CC occurring enzyme, but has higher thermostability. Cholesterol
 CC oxidase DNA sequences derived from Streptomyces sp. are given in
 CC AAQ70417-18.
 XX
 SQ Sequence 1575 BP; 281 A; 599 C; 478 G; 217 T; 0 other;

Query Match 6.6%; Score 39.6; DB 15; Length 1575;
 Best Local Similarity 47.2%; Pred. No. 0.086;
 Matches 120; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
 OY 212 accctgttcgaagatcgcccaagatccaactgaaccgaagctcaagccgagctgtg 271
 DB 1217 acgacgcgcgaagcagcgcgcgaagctcaactgaaccgctgacccaagcccccgcg 1276
 OY 272 ttaacgcttcacagatgtgttcgtcgaagctatgaagtcggcgaagcctgaagctca 331
 DB 1277 tcaacgacgaagcagcgtgttcgacgcgatacaagaagcgaagcagcgtatccggt 1336
 OY 332 ccgagctgttctcgcgtcgaagcgcgtcaagccgcgcgtcgaaccgcgcgaaccgcgca 391
 DB 1337 acgaaccttcgcgcacccagctgaagccttcgcgcgaagccttcgtaccaccgcgtcg 1396
 OY 392 ctggcgaagcagatgacatccttcgcgcgttcctgatactccgctgtcctcctgc 451
 DB 1397 gcggcgcgcgtcctcgcgcgaagcgcgcgaagcgcgcgcgcgcgttcgaagaacc 1456
 OY 452 tgaagaagcgcgtc 465
 DB 1457 tctaagtgaccgcac 1470

RESULT 11

AAQ70418
 ID AAQ70418 standard; DNA; 1578 BP.
 XX
 AC AAQ70418;
 XX
 DT 01-MAR-1995 (first entry)
 XX
 DE Streptomyces cholesterol oxidase.
 XX
 KW Cholesterol oxidase; Streptomyces; enzyme; thermostability;
 XX diagnosis; endocrine disease; metabolism; ss.
 XX
 OS Streptomyces sp. SA-COO.
 XX
 PN JP06169765-A.
 XX
 PD 21-JUN-1994.
 XX
 PF 05-JUL-1993; 93JP-0165558.
 XX
 PR 05-OCT-1992; 92JP-0266130.
 XX
 PA (TOYM) TOYBO KK.
 XX
 XX WPI: 1994-237584/29.
 DR P-PSDB; AAR60325.
 XX
 PT Cholesterol oxidase with high thermostability - useful for the
 XX diagnosis of endocrine diseases
 PS Claim 8; Page 11-13; 15pp; Japanese.

XX
 CC Cholesterol oxidase is used for measurement of cholesterol which is
 CC an index for the diagnosis of endocrine diseases and abnormal
 CC metabolism. The enzyme is obtained by adding amino acid residues to
 CC the N-terminus of cholesterol oxidase derived from Streptomyces sp.
 CC SA-COO, which shows the same activity as that of the naturally
 CC occurring enzyme, but has higher thermostability. Cholesterol
 CC oxidase DNA sequences derived from Streptomyces sp. are given in
 CC AAQ70417-18.
 XX
 SQ Sequence 1578 BP; 281 A; 601 C; 479 G; 217 T; 0 other;

Query Match 6.6%; Score 39.6; DB 15; Length 1578;
 Best Local Similarity 47.2%; Pred. No. 0.086;
 Matches 120; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
 OY 212 accctgttcgaagatcgcccaagatccaactgaaccgaagctcaagccgagctgtg 271
 DB 1220 acgacgcgcgaagcagcgcgcgaagctcaactgaaccgctgacccaagcccccgcg 1279
 OY 272 ttaacgcttcacagatgtgttcgtcgaagctatgaagtcggcgaagcctgaagctca 331
 DB 1280 tcaacgacgaagcagcgtgttcgacgcgatacaagaagcgaagcagcgtatccggt 1339
 OY 332 ccgagctgttctcgcgtcgaagcgcgtcaagccgcgcgtcgaaccgcgcgaaccgcgca 391
 DB 1340 acgaaccttcgcgcacccagctgaagccttcgcgcgaagccttcgtaccaccgcgtcg 1399
 OY 392 ctggcgaagcagatgacatccttcgcgcgttcctgatactccgctgtcctcctgc 451
 DB 1400 gcggcgcgcgtcctcgcgcgaagcgcgcgaagcgcgcgcgcgcgttcgaagaacc 1459
 OY 452 tgaagaagcgcgtc 465
 DB 1460 tctaagtgaccgcac 1473

RESULT 12

AAT49320
 ID AAT49320 standard; DNA; 1638 BP.
 XX
 AC AAT49320;
 XX
 DT 18-FEB-1997 (first entry)
 XX
 DE Cholesterol oxidase coding sequence.
 XX
 KW Cholesterol oxidase; mutation; thermal stability; substitution; ss.
 XX
 OS Streptomyces sp. SA-COO.
 XX
 PN JP06242860-A.
 XX
 PD 24-SEP-1996.
 XX
 PF 07-MAR-1995; 95JP-0047339.
 XX
 PR 07-MAR-1995; 95JP-0047339.
 XX
 PA (TOYM) TOYBO KK.
 XX
 XX WPI: 1996-479904/48.
 DR P-PSDB; AAM08930.
 XX
 PT Modified cholesterol oxidase enzyme - having higher thermal
 XX stability than the parent enzyme, useful in cholesterol determin.
 PS Disclosure: Page 19-21; 23pp; Japanese.
 CC This sequence encodes wild type cholesterol oxidase. Modified
 CC cholesterol oxidases may be prepared by mutation of this

PS Disclosure: Page 100-102; 187pp; Japanese.

CC The present invention describes cells originating in bone marrow or
 CC umbilical blood cells which are capable of differentiating into
 CC cardiomyocytes. Also described are: (1) cardiomyocytes produced by the
 CC differentiation of the cells; (2) a method for carrying out the
 CC differentiation into cardiomyocytes, regulated by a promotional and/or
 CC inhibitory factor; (3) a method for the differentiation of the cells
 CC into cell types other than cardiomyocytes; (4) drug compositions
 CC promoting the formation of heart muscle and regeneration of heart tissue
 CC which contain the cells; (5) a method for the production of antibodies
 CC which recognise the cells; (6) a method for screening factors which
 CC promote antigen on the cells; (7) a method for screening factors which
 CC the cells by expressing telomerase in them; (8) drug compositions for
 CC the treatment of heart disease which contain the immortalised cells; and
 CC (9) cell-free supernatant from the culture of the cells and its use in
 CC promoting their differentiation into cardiomyocytes. The cells are used
 CC in the treatment of diseases involving heart muscle degeneration, such
 CC as myocardial infarction and in the study of cardiomyocyte
 CC differentiation. AAH44351 to AAH44409 and AAB99935 represent
 CC sequences used in the exemplification of the present invention.

XX Sequence 1326 BP; 215 A; 499 C; 414 G; 198 T; 0 other;

Query Match 6.5%; Score 39; DB 22; Length 1326;

Best Local Similarity 56.4%; Pred. No. 0.12;
 Matches 92; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

QY 235 aagtcacactgacacaggtcagcagcgtctgttaacgcttcacagatgtgttc 294

DB 566 AGGACCGGGCTGTGCAAGGGGCGCGGAGCGGCGGCTGCGGCGGCGGCGCC 507

QY 295 gtcgaggtatgaagtcgagcagcgtctgaagtcacacgctgttcctcgtagcgc 354

DB 506 ACGTCGGCCATGTAAAGCCGGGTAGGGGCTGGAG-TAGGAGCCCGGAGACCGCGCGCC 448

QY 355 gtcaagcgccggtcgcagcagcgcacacccgcgcactgagcg 397

DB 447 GTACTGCTCGCGCGCGCCGACGCGCCGACCCGCTCGCGCG 405

RESULT 15

AAH48225/c

ID AAH48225 standard; DNA; 1326 BP.

AC AAH48225;

DT 21-SEP-2001 (first entry)

DE Heart muscle cell differentiation related DNA SEQ ID NO: 12.

KW Heart muscle cell; human; cell differentiation; heart disease; ds.

OS Homo sapiens.

PN WO200148151-A1.

PD 05-JUL-2001.

PF 27-DEC-2000; 2000WO-JP09323.

PR 28-DEC-1999; 99JP-0372826.

PR 28-FEB-2000; 2000WO-JP01148.

PR 02-NOV-2000; 2000WO-JP07741.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;

XX Yamada Y;

XX WPI: 2001-425656/45.

DR P-PSDB; AAG64849.

XX Cells capable of differentiating into cardiomyocytes and originating in
 PT bone marrow or umbilical blood cells for study of cardiomyocyte
 PT differentiation and treatment of heart disease

PS Disclosure: Page 106-108; 183pp; Japanese.

CC The present invention provides cells originating in the human bone marrow
 CC or umbilical blood cells which are capable of differentiating into
 CC cardiomyocytes. These cells are useful in the treatment of diseases
 CC involving heart muscle degeneration, such as myocardial infarction, and
 CC the study of cardiomyocyte differentiation. The present sequence is
 CC an oligonucleotide described in the exemplification of the invention.

XX Sequence 1326 BP; 215 A; 499 C; 414 G; 198 T; 0 other;

Query Match 6.5%; Score 39; DB 22; Length 1326;

Best Local Similarity 56.4%; Pred. No. 0.12;
 Matches 92; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

QY 235 aagtcacactgacacaggtcagcagcgtctgttaacgcttcacagatgtgttc 294

DB 566 AGGACCGGGCTGTGCAAGGGGCGCGGAGCGGCGGCTGCGGCGGCGGCGCC 507

QY 295 gtcgaggtatgaagtcgagcagcgtctgaagtcacacgctgttcctcgtagcgc 354

DB 506 ACGTCGGCCATGTAAAGCCGGGTAGGGGCTGGAG-TAGGAGCCCGGAGACCGCGCGCC 448

QY 355 gtcaagcgccggtcgcagcagcgcacacccgcgcactgagcg 397

DB 447 GTACTGCTCGCGCGCGCCGACGCGCCGACCCGCTCGCGCG 405

Search completed: December 20, 2001, 20:02:58
 Job time: 4094 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2001, 18:51:34 ; Search time 88.01 seconds
(without alignments)
1543.993 Million cell updates/sec

Title: US-09-816-391A-1

Perfect score: 600
1 gctggcgccgagccatga.....tattccgactagtcagcg 600

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents: NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCNUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Bactfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43.6	7.3	16075	3	US-09-096-942-1
2	43.6	7.3	16075	3	US-09-096-867-1
3	37.6	6.3	858	4	US-09-095-855-200
4	37	6.2	2943	1	US-08-042-747A-7
5	35	5.8	1950	3	US-08-911-853-14
6	35	5.8	1950	4	US-09-479-409-14
7	35	5.8	17612	3	US-08-911-853-29
8	35	5.8	17612	4	US-09-479-409-29
9	35	5.8	4403765	4	US-09-103-840A-2
10	35	5.8	4411529	4	US-09-103-840A-1
11	34.8	5.8	2721	6	5215881-2
12	34.8	5.8	8438	1	US-07-945-283-1
13	34.6	5.7	1524	2	US-08-426-598B-3
14	34.4	5.7	1524	6	5512669-1
15	34.4	5.7	6854	4	5229279-1
16	34	5.7	8323	1	US-09-194-905-7
17	34	5.7	8323	2	US-08-110-300A-8
18	34	5.7	8323	5	US-08-886-642-8
19	34	5.7	8323	8	US-08-886-642-8
20	34	5.7	10367	2	PCT-US93-08041-8
21	34	5.7	10367	1	US-08-110-300A-9
22	34	5.7	10367	5	PCT-US93-08041-9
23	33.8	5.6	1710	2	US-08-912-794-1
24	33.8	5.6	6379	1	US-08-489-215-1
25	33.2	5.5	337	2	US-08-997-080-191
26	33.2	5.5	337	2	US-08-997-362-191
27	33.2	5.5	337	4	US-09-095-855-191

C	28	33	5.5	80161	3	US-09-036-987A-1	Sequence 1, Appli
C	29	33	5.5	891	4	US-09-370-700-1	Sequence 1, Appli
C	30	32.8	5.5	12588	2	US-09-049-672A-23	Sequence 23, Appli
C	31	32.8	5.5	1521	1	US-08-387-942C-1	Sequence 1, Appli
C	32	32.4	5.4	1521	1	US-08-083-948-14	Sequence 14, Appli
C	33	32.4	5.4	1521	1	US-08-393-785-14	Sequence 14, Appli
C	34	32.4	5.4	1521	1	US-08-475-694-14	Sequence 14, Appli
C	35	32.4	5.4	1521	1	US-08-712-057-14	Sequence 14, Appli
C	36	32.4	5.4	1647	1	US-08-083-948-13	Sequence 13, Appli
C	37	32.4	5.4	1647	1	US-08-393-785-13	Sequence 13, Appli
C	38	32.4	5.4	1647	1	US-08-475-694-13	Sequence 13, Appli
C	39	32.4	5.4	1647	1	US-08-712-057-13	Sequence 13, Appli
C	40	32.4	5.4	1865	1	US-08-083-948-7	Sequence 7, Appli
C	41	32.4	5.4	1865	1	US-08-393-785-7	Sequence 7, Appli
C	42	32.4	5.4	1865	1	US-08-475-694-7	Sequence 7, Appli
C	43	32.4	5.4	1865	1	US-08-712-057-7	Sequence 7, Appli
C	44	32.2	5.4	15872	4	US-09-105-537-1	Sequence 1, Appli
C	45	32.2	5.4	36741	4	US-09-301-665-3	Sequence 3, Appli

ALIGNMENTS

```

RESULT 1
US-09-096-942-1
: Sequence 1, Application US/09096942
: Patent No. 6027925
: GENERAL INFORMATION:
: APPLICANT: Pollock, Thomas J
: APPLICANT: Mikolajczak, Marcia
: APPLICANT: Yamazaki, Motohide
: APPLICANT: Thorne, Linda
: APPLICANT: Armentrout, Richard W
: TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
: TITLE OF INVENTION: Carrying Genes from Xanthomonas Campestris
: FILE REFERENCE: seq list for appli filed from pro. appli
: CURRENT APPLICATION NUMBER: US/09/096,942
: EARLIER FILING DATE: 1998-06-12
: EARLIER APPLICATION NUMBER: 60/049,428
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 1
: LENGTH: 16075
: TYPE: DNA
: ORGANISM: Xanthomonas campestris
: PUBLICATION INFORMATION:
: US-09-096-942-1

```

Query Match 7.3%, Score 43.6; DB 3; Length 16075;
Best Local Similarity 47.7%; Pred. No. 0.0017;
Matches 127; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

OY	123	atggatacaacagctcgtactcgttcgaaagatcgcgcgaagttccaactgcgaag	252
DB	33	atggatctgaagcaagcggaagatgctcgagcgctcgttcgacgaagttcgctgaacag	92
OY	253	gtccagcgccagcgctcgttcgaacgccttccaagatgcttcgttcgagcgatgaagcc	312
DB	93	cgtagggaggaaggaattcgtcgaacgcgttttcgatgtgctcgcgatgaatgaagcaag	152
OY	313	ggcgaagcctgaagctacccgctgttcctcgctgagcgcgltcaagcgccgctcgc	372
DB	153	ggccctcaggtgaagttgtcggtcgttcgcaacttcgtcgcgcgcaagaaccaacg	212
OY	373	acggcgccgaacccgagcgcgacgagcgagcagatgtgacatccggtccttaagcgctt	432
DB	213	ccggctcgaatcccaagacccggtgaagaaatccgaltcgcgcgaagcagcgtgtgacc	272
OY	433	atctcgcgtcctcctctgtgaagaa	458
DB	273	ttccgccccgcgcgaactcaagaa	298

RESULT 2

US-09-096-867-1

Sequence 1, Application US/09096867

Patent No. 6030817

GENERAL INFORMATION:

APPLICANT: Pollock, Thomas J

APPLICANT: Mikolajczak, Marcia

APPLICANT: Yamazaki, Motohide

APPLICANT: Thorne, Linda

APPLICANT: Armentrout, Richard W

TITLE OF INVENTION: Production of Xanthan Gum by *Sphingomonas Bacteria*TITLE OF INVENTION: Carrying Genes from *Xanthomonas Campestris*

FILE REFERENCE: seq list for appl filed from pro. appl

CURRENT APPLICATION NUMBER: US/09/096,867

EARLIER FILING DATE: 1998-06-11

EARLIER APPLICATION NUMBER: 60/049,428

NUMBER OF SEQ ID NOS: 2

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 1

LENGTH: 16075

TYPE: DNA

ORGANISM: *Xanthomonas campestris*

PUBLICATION INFORMATION:

US-09-096-867-1

Query Match 7.3%, Score 43.6; DB 3; Length 16075;

Best Local Similarity 47.7%; Pred. No. 0.0017;

Matches 127; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

```
OY 193 atggatatacaacagctgccttcggaagatcgcccaagaatcccaactgacaaag 252
    ||||| || || || || || || || || || || || || || || || || || ||
DB 33 atgcatgaagaagcggagatgagccgagcgtctgttcgaagaagtcggcctgaacaa 92
    ||||| || || || || || || || || || || || || || || || || || ||
OY 253 gctcagcgcggagctgctgttaacgcctccagagatggtgtcgaagctgaagttcc 312
    ||||| || || || || || || || || || || || || || || || || || ||
DB 93 cgtgagggaggaattcgtcgaacggttttcgattgtgtcgcgactgagagcag 152
    ||||| || || || || || || || || || || || || || || || || || ||
OY 313 ggcgaagcctgaagctcaacgcgtgtttctcgtgagcggtcaagcgcccgctgcg 372
    ||||| || || || || || || || || || || || || || || || || || ||
DB 153 ggcgcagcgtgaagtgtcgtgcttcgcgaacttcgactcgcgcgcaagaacacgg 212
    ||||| || || || || || || || || || || || || || || || || || ||
OY 373 accgcgcaacccgcgactgctgcgagcagattgacattccgcttcctacgcgttcgt 432
    ||||| || || || || || || || || || || || || || || || || || ||
DB 213 ccggtcgaatcccaagacgcggtgaggaattccgactcgtcgcgacgagcggtgtgacc 272
    ||||| || || || || || || || || || || || || || || || || || ||
OY 433 atctcgcgtgctccctcgtgaagaa 458
    ||||| || || || || || || || || || || || || || || || || || ||
DB 273 ttccgccccgcgcagaaactcaagaa 298
    ||||| || || || || || || || || || || || || || || || || || ||
```

RESULT 3

US-09-095-855-200

Sequence 200, Application US/09095855

Patent No. 6160093

GENERAL INFORMATION:

APPLICANT: Tan, Paul

APPLICANT: Visser, Elizabeth

APPLICANT: Skinner, Margot

APPLICANT: Prestidge, Ross

TITLE OF INVENTION: Compounds and Methods for *Mycobacterial Infections*TITLE OF INVENTION: Treatment and Diagnosis of *Mycobacterial Infections*

NUMBER OF SEQUENCES: 208

CORRESPONDENCE ADDRESS:

ADDRESS: Law Offices of Ann W. Speckman

STREET: 2601 Elliott Avenue, Suite 4185

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/095,855

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/705,347

FILING DATE: 29-AUG-1996

APPLICATION NUMBER: 08/873,970

FILING DATE: 12-JUN-1997

APPLICATION NUMBER: 08/997,362

FILING DATE: 23-DEC-1997

ATTORNEY/AGENT INFORMATION:

NAME: Sleath, Janet

REGISTRATION NUMBER: 37,007

REFERENCE/DOCKET NUMBER: 11000.1002c3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-269-0565

TELEFAX: 206-269-0563

TELEX:

INFORMATION FOR SEQ ID NO: 200:

SEQUENCE CHARACTERISTICS:

LENGTH: 858 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-09-095-855-200

Query Match 6.3%, Score 37.6; DB 4; Length 858;

Best Local Similarity 46.3%; Pred. No. 0.027;

Matches 124; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

```
OY 202 aacaaatcgtcctcgtttcgaagatcgcccaagaatcccaactgacaaagctcaagcc 261
    ||||| || || || || || || || || || || || || || || || || || ||
DB 187 AACAAAGCAGAGCTCATGACGTACTGACGAGAGCTGCGCTCGATCGATCGGCAAGCG 246
    ||||| || || || || || || || || || || || || || || || || || ||
OY 262 gagctgtgttaagcttcacagatgtgttcgtcgaagtgatataagtcggcgaaagc 321
    ||||| || || || || || || || || || || || || || || || || || ||
DB 247 ACTGCGCGGTGGAAGAGTGTCTGACACCATCGTGCCTGCTGCTGCTGCTGCTGCTGCT 306
    ||||| || || || || || || || || || || || || || || || || || ||
OY 322 ctgaagctcacgcgctgttcctcgtcgtcgaagcgcgcgcgcgcgcgcgcgcgcgcgcgc 381
    ||||| || || || || || || || || || || || || || || || || || ||
DB 307 GTCACCATCAGCGGCTTGGGTGTTTTCGAGCAGCGTCTGCGGCGACGACGCGTGCACAGC 366
    ||||| || || || || || || || || || || || || || || || || || ||
OY 382 aaccgcgacactggcgcagagattgacattccgacttcctcgaagcgttcgtatctcgct 441
    ||||| || || || || || || || || || || || || || || || || || ||
DB 367 AATCCGCGCAGCGGCGGAGACCGTGAAGGTCAAGCCACCTCAGTCCGCGATTCGCTCC 426
    ||||| || || || || || || || || || || || || || || || || || ||
OY 442 ggtcctcgtcgaagaagcgctcacgcg 469
    ||||| || || || || || || || || || || || || || || || || || ||
DB 427 GCGCTCAGTTCAAGGCTGTGTCTCTG 454
    ||||| || || || || || || || || || || || || || || || || || ||
```

RESULT 4

US-08-042-747A-7

Sequence 7, Application US/08042747A

Patent No. 5487969

GENERAL INFORMATION:

APPLICANT: Eberle, Richard

APPLICANT: Black, Darla

APPLICANT: Scintaricello, Franco

APPLICANT: Hilliard, Julia K.

TITLE OF INVENTION: Cloning and Amplification of Monkey B

TITLE OF INVENTION: Virus Genes

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESS: Cox & Smith Incorporated


```

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtis P
REGISTRATION NUMBER: 27976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 309-685-4128
TELEFAX: 309-685-4011 ext.513
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8438 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOHETICAL: NO
AMTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pseudorabies virus
FEATURE:
NAME/KEY: CDS
LOCATION: 622..6495
FEATURE:
NAME/KEY: variation
LOCATION: replace(1099, "g")
FEATURE:
NAME/KEY: variation
LOCATION: replace(1267, "t")
FEATURE:
NAME/KEY: variation
LOCATION: replace(1381, "c")
FEATURE:
NAME/KEY: variation
LOCATION: replace(1566, "c")
FEATURE:
NAME/KEY: variation
LOCATION: replace(7010, "g")
US-07-945-283-1

```

```

Query Match          5.8%; Score 34.8; DB 1; Length 8438;
Best Local Similarity 54.8%; Pred. No. 0.68;
Matches 69; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 315 cgaagcctgaagctaacgcgctgtctcgtgtagcgctgaagcgcgcgcgtcgac 374
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 3583 CTACGACCCGCTCCCGACGAGAGCTCTCGCCGACCCCGAGGGCGCCCGCGCCC 3524

QY 375 cgcgcgaacccgcgcactggcgagcagatgaattccggtcttcctcagggcttgat 434
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 3523 GGCCTGGAGACCGCGCGCGCGCGCTCTCTGGGCGCTCTCGCCGCTTCGCCACCGCCT 3464

QY 435 ctcgcgc 440
    |||||
DB 3463 CTCGAC 3458

RESULT 13
US-08-426-599B-3
; Sequence 3, Application US/08426599B
; Patent No. 5932438
; GENERAL INFORMATION:
; APPLICANT: Uriach-Marsal, Juan
; APPLICANT: Rubio-Susan, Victor
; APPLICANT: Palino-Martín, Cristina
; APPLICANT: Kalo-Koenova, Eliza I.
; APPLICANT: del Moral-Juarez, Catalina
; APPLICANT: Faus-Santasusana, Ignacio
; APPLICANT: del Río-Pericacho, Jose-Luis
; APPLICANT: Blade-Plaque, Joan

```

```

TITLE OF INVENTION: Preparation of Thaumatin Sweeteners
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
STREET: 555 Thirteenth Street, N.W., Suite 701 East
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,599B
FILING DATE: 21-APR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Stephen A.
REGISTRATION NUMBER: 38,609
REFERENCE/DOCKET NUMBER: 1604-123A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-624-1589
TELEFAX: 202-783-6031
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Optimized cDNA"
HYPOHETICAL: NO
AMTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..621
US-08-426-599B-3

```

```

Query Match          5.8%; Score 34.6; DB 2; Length 624;
Best Local Similarity 50.9%; Pred. No. 0.2;
Matches 82; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 224 agtgcgccagaagtcacactgaccagagctcagcgccgaggtgtgttaagccttc 283
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 11 AGATCGTCAACCGCTCTCTCTACACCGTCTGGCGCCCGCTTCAAGGGAGCGCGCCC 70

QY 284 aggatgttgcgcgagggcgtatgaagtcgcggaagcctgaagctcacggcgttct 343
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 71 TCGACCGCGCGCGCGCGCGCGCTCACTCGGCGAGTCTGGACCATCAACGTGAGACCG 130

QY 344 cgcgtgagcgcgtcaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 384
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 131 GCACCAAGGGCGGCAAGATCTGGGCGCGCGCGCGCGCGCTGCTAC 171

RESULT 14
5512669-1
; Patent No. 5512669
; APPLICANT: PEOPLES, OLIVER P.; SINSKEY, ANTHONY J.
; TITLE OF INVENTION: GENE ENCODING BACTERIAL ACETOACETYL-COA
; REDUCTASE
; NUMBER OF SEQUENCES: 7
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/297,667
; FILING DATE: 29-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 124,570
; FILING DATE: 20-SEP-1993
; APPLICATION NUMBER: 944,488

```

FILED DATE: 03-NOV-1992
APPLICATION NUMBER: 566,535
FILED DATE: 13-AUG-1990
APPLICATION NUMBER: 67,695
FILED DATE: 29-JUN-1987
SEQ ID NO:1:
LENGTH: 1524
5512669-1

Query Match 5.7%; Score 34.4; DB 6; Length 1524;
Best Local Similarity 49.4%; Pred. No. 0.37;
Matches 89; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

OY 201 caacaagctgaccctcgttcgaagatcgcccaagaatcccaactgaccgaagctcagc 260
DB 359 caacagcgcgcgcacatgaactcggggcgacgctgattcgcggtctcgaagcgcgg 418
OY 261 cgaagctgctgttaacgcttcacagatgtgttcgtagcgtatgaagtcgcgcgaag 320
DB 419 cgttcggcgcgagcgagtgatgtatcttcgccaagtgctgcgcgcgcgcgaag 478
OY 321 cctgaagctcacggcctgttctccgtgagcgctcaagcgccgcgctcgaccgcgcg 380
DB 479 ccagaccgcgcgcgcacatgaagcgcgctgctccgcaggaagcgaccgcctg 538

RESULT 15

5229279-1
PATENT NO. 5229279
APPLICANT: PEOPLES, OLIVER P.; SINSKEY, ANTHONY J.
TITLE OF INVENTION: METHOD FOR PRODUCING NOVEL POLYESTER
BIOPOLYMERS
NUMBER OF SEQUENCES: 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/556,535
FILING DATE: 13-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 67,695
FILING DATE: 29-AUG-1987
SEQ ID NO:1:
LENGTH: 1525
5229279-1

Query Match 5.7%; Score 34.4; DB 6; Length 1525;
Best Local Similarity 49.4%; Pred. No. 0.37;
Matches 89; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

OY 201 caacaagctgaccctcgttcgaagatcgcccaagaatcccaactgaccgaagctcagc 260
DB 360 caacagcgcgcgcacatgaactcggggcgacgctgattcgcggtctcgaagcgcgg 419
OY 261 cgaagctgctgttaacgcttcacagatgtgttcgtagcgtatgaagtcgcgcgaag 320
DB 420 cgttcggcgcgagcgagtgatgtatcttcgccaagtgctgcgcgcgcgcgaag 479
OY 321 cctgaagctcacggcctgttctccgtgagcgctcaagcgccgcgctcgaccgcgcg 380
DB 480 ccagaccgcgcgcgcacatgaagcgcgctgctccgcaggaagcgaccgcctg 539

Search completed: December 20, 2001, 20:35:15
Job time: 6221 sec

OY 384 cccgcgcactgagcgagcagatgacattccggtctctacgagcggttcgattccgctg 443
 DB 133 CCCGACGACCGGCGGCTCCAGATCCCGCCAGCAGGCGCCGCTTCAAGGCTTC 242
 OY 444 ctccctgtcgaagaagcgctcacg 469
 DB 243 AGTTGGCTGAAGATGCCGTGAACG 268

RESULT 4

AV387318 509 bp mRNA EST 29-SEP-2000
 LOCUS AV387318 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
 DEFINITION CDNA clone CM019f02_r, mRNA sequence.
 ACCESSION AV387318
 VERSION AV387318.1 GI:6541534
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii.
 ORGANISM Chlamydomonas reinhardtii.
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadae; Chlamydomonas.

REFERENCE 1 (bases 1 to 509)
 ASAMIZU, E., NAKAMURA, Y., SATO, S., FUKUZAWA, H. and TABATA, S.
 A large scale structural analysis of cDNAs in a unicellular green
 non-redundant expressed sequence tags

JOURNAL DNA Res. 6 (6), 369-373 (1999)
 MEDLINE 20152988

COMMENT

Contact: Yasukazu Nakamura
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 153-3, Kisarazu, Chiba 292-0812, Japan
 Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 Location/Qualifiers
 1. 509
 /organism="Chlamydomonas reinhardtii"
 /strain="C9"
 /db_xref="taxon:3055"
 /clone="CM019f02_r"
 /clone_1lb="Chlamydomonas reinhardtii C9"
 /dev_stage="photoautotrophic growth"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 82 a 182 c 168 g 77 t
 ORIGIN

Query Match 9.9%; Score 59.6; DB 10; Length 509;
 Best Local Similarity 51.5%; Pred. No. 1.2e-06;
 Matches 137; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

OY 204 caagctgacctgcttgcgaagatcgccagaagtcacactgacgaagctcagcgca 263
 DB 125 CAAGGCTAAACTTGTGAGAGCGATCGCTACTGAGCTGGAGTGAACAGACGTCGTGC 184
 OY 264 gctctgtttaacgcttccagaatgtgttcgtcgaagctatgaagtcgagcgca 323
 DB 185 GAAGGCTTTGACTCCCTGATTTGGCGGCAATTGAGAGCGCCCTCATCAACGGTGAACCCGCT 244
 OY 324 gaagctcaccgagcgtgttctcgcgtcgaagcgctcaagcgccgagctcgacacgagcgca 383
 DB 245 GACCATCTGTGGCTTGGGACGTTTGAAGTGCAGCGGCGCTGCGCCAGGCGCCGCA 304
 OY 384 ccgcgcactgagcgagcagatgacattccggtctctacgagcggttcgattcccgctg 443
 DB 305 CCCGACGACCGGCGGCTCCAGATCCCGCCAGCAGGCGCCCGCTTCAAGGCTTC 364
 OY 444 ctccctgtcgaagaagcgctcacg 469
 DB 365 AGTTGGCTGAAGATGCCGTGAACG 390

RESULT 5

BE337777 625 bp mRNA EST 14-JUL-2000
 LOCUS BE337777 894048D02.y1 C. reinhardtii CC-1690, normalized, lambda Zap II
 DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.
 ACCESSION BE337777
 VERSION BE337777.1 GI:9210862
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii.
 ORGANISM Chlamydomonas reinhardtii.
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadae; Chlamydomonas.

REFERENCE 1 (bases 1 to 625)
 GROSSMAN, A., DAVIES, J., FEDERSPIEL, N., HARRIS, E., LEFEBVRE, P.,
 MCDERMOTT, J. P., SILFLOW, C., STERN, D. and SURZYCKI, R.

Analyses of the Chlamydomonas reinhardtii Genome: A Model,
 Unicellular System for Analyzing Gene Function and Regulation in
 Vascular Plants; Project phase 2
 Unpublished (2000)

JOURNAL

Contact: Elizabeth H. Harris
 Duke University
 Durham, NC 27708-1000, USA
 Tel.: 919 613 8164
 Fax: 919 613 8177
 Email: chlamy@duke.edu.

FEATURES

SOURCE

1. 625
 /organism="Chlamydomonas reinhardtii"
 /strain="CC-1690 wild type mt+ 21gr"
 /db_xref="taxon:3055"
 /clone_1lb="C. reinhardtii CC-1690, normalized, Lambda Zap
 II"

/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
 XhoI; This library, constructed by John Davies and Jeffrey
 McDermott, combines cDNAs from CC-1690 cells grown to
 mid-log phase in TAP (acetate-containing) medium in the
 light, TAP medium in the dark, HS (minimal) medium in
 ambient levels of CO2 and HS medium bubbled with 5% CO2.
 PolyA mRNA was purified from each sample, pooled and cDNA
 synthesized. The cDNA was directionally cloned into lambda
 Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
 pBluescript II SK- plasmids were excised from the lambda
 Zap clones by superinfection with Exsist (Stratagene)
 phage. The library was normalized using method 4 described
 in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT 106 a 216 c 198 g 104 t
 ORIGIN

Query Match 9.9%; Score 59.6; DB 10; Length 625;
 Best Local Similarity 51.5%; Pred. No. 1.2e-06;
 Matches 137; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

OY 204 caagctgacctgcttgcgaagatcgccagaagtcacactgacgaagctcagcgca 263
 DB 154 CAAGGCTAAACTTGTGAGAGCGATCGCTACTGAGCTGGAGTGAACAGACGTCGTGC 213
 OY 264 gctctgtttaacgcttccagaatgtgttcgtcgaagctatgaagtcgagcgca 323
 DB 214 GAAGGCTTTGACTCCCTGATTTGGCGGCAATTGAGAGCGCCCTCATCAACGGTGAACCCGCT 273
 OY 324 gaagctcaccgagcgtgttctcgcgtcgaagcgctcaagcgccgagctcgacacgagcgca 383
 DB 274 GACCATCTGTGGCTTGGGACGTTTGAAGTGCAGCGGCGCTGCGCCAGGCGCCGCA 333
 OY 384 ccgcgcactgagcgagcagatgacattccggtctctacgagcggttcgattcccgctg 443
 DB 334 CCCGACGACCGGCGGCTCCAGATCCCGCCAGCAGGCGCCCGCTTCAAGGCTTC 393
 OY 444 ctccctgtcgaagaagcgctcacg 469
 DB 394 AGTTGGCTGAAGATGCCGTGAACG 419

RESULT	6
LOCUS	BG860650
DEFINITION	BG860650 696 bp mRNA EST 29-MAY-2001
ACCESSION	1024072A08.y1 C. reinhardtii CC-1690, normalized, Lambda Zap II
VERSION	BG860650
KEYWORDS	BG860650.1 GI:14241834
SOURCE	EST.
ORGANISM	Chlamydomonas reinhardtii.
REFERENCE	Chlamydomonas reinhardtii.
AUTHORS	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.
TITLE	1 (bases 1 to 696)
JOURNAL	Grossman, A., Davies, J., Pederspiel, N., Harris, E., Lefebvre, P.,
COMMENT	Modernot, J. P., Sillflow, C., Stern, D. and Surzycki, R.
	Analyses of the Chlamydomonas reinhardtii Genome: A Model,
	Unicellular System for Analyzing Gene Function and Regulation in
	Vascular Plants; project phase 2
	Unpublished (2000)
	Contact: Charles Hauser
	DCMB Box 91000
	Duke University
	Durham, NC 27708-1000
	Tel: 919 613 8159
	Fax: 919 613 8177
	Email: chauser@duke.edu.
FEATURES	location/Qualifiers
SOURCE	1. 696
	/organism="Chlamydomonas reinhardtii"
	/strain="CC-1690 wild type mt+ 21gr"
	/db_xref="taxon:3055"
	/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
	II"
	/note="Vector: pBluescript II SK-; Site.1: EcoRI; Site.2:
	XhoI; This library, constructed by John Davies and Jeffrey
	McMurtro, combines cDNAs from CC-1690 cells grown to
	mid-log phase in YAP (acetate-containing) medium in the
	light, YAP medium in the dark, HS (minimal) medium in
	ambient levels of CO2 and HS medium bubbled with 5% CO2.
	PolyA mRNA was purified from each sample, pooled and cDNA
	synthesized. The cDNA was directionally cloned into lambda
	Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
	pBluescript II SK- plasmids were excised from the lambda
	Zap clones by superinfection with Exassist (Stratagene)
	phage. The library was normalized using method 4 described
	in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT	120 a 238 c 222 g 116 t
ORIGIN	

Query Match	9.9%	Score 59.6	DB 11	Length 696
Best Local Similarity	51.5%	Pred. No. 1.3e-06		
Matches 137	Conservative 0	Mismatches 129	Indels 0	Gaps 0
QY 204	caagctcyacctcgcttctcgaagaatctcgcccaagaagtcacaacttgaacaaagctcagcgoga	263		
DB 137	CAAGGCTTAAACTGTGGAGGCGATGCTACTAGGNGGCACTGACCAAGACGTCGCTCC	196		
QY 264	gcctcgtgtlaacgctcccaagatcgtgttcgcgaagctatgaaatccgcgcgaagcct	323		
DB 197	GAAGGCCCTTTGACCTCCCGATGTTGGGGCGCATTTAGAGAGCCCTCATCAACGGTGAACGGCT	256		
QY 324	gaagctcaacgcgcctgtctctccgcctgtagcgcgctcaagcgccgcgcctcgcaaccgcgcgcaa	383		
DB 257	GACCATCTCGTGGCGCTTCGGACACGCTTTGAGGTGGCGCGCCGCTGGCGCGACAGGCGCGCA	316		
QY 384	cccgcgcaactgcgcgaagacatgacatccgccttcctcactcagcgcyttcgtatctccgcctg	443		
DB 317	CCCCAGCAACGGGGGGGTCCTCCAGATGCCGCCCAAGAGGCCGCCCGTTCAAAGGCTTC	376		
QY 444	ctccctcgtgaagaagccgctcacgg	469		

Db	377	AGTGGCCTGAAGATGCGTGACG	A02
RESULT	7		
LOCUS	AQ990760/c		
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
COMMENT			

AQ990760 592 bp DNA GSS 14-AUG-2000
 RC01581 Photorhabdus luminescens strain W14 M13 library
 Photorhabdus luminescens genomic clone pLG01581, DNA sequence.
 AQ990760
 AQ990760.1 GI:9649354
 GSS.
 Photorhabdus luminescens.
 Photorhabdus luminescens
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Photorhabdus.
 1 (bases 1 to 592)
 french-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,
 Daborn,P.J., Bowen,D. and Blattner,F.R.
 A genomic sample sequence of the entomopathogenic bacterium
 Photorhabdus luminescens W14: potential implications for virulence
 Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)
 20378633
 Contact: french-Constant RH
 Department of Biology and Biochemistry
 University of Bath
 South Building, Bath BA2 7AY, UK
 Tel: (44) 1225 826621
 Fax: (44) 1225 826779
 Email: bsr@cepath.ac.uk
 This is one of 2,122 random reads from the M13 library. For
 annotation of identified clones (BLASTX, BLASTN and mapping to E.
 coli K12 genome) please see french-Constant et al. 2000, Nucleic
 Acids Res.
 Seq primer: M13 Forward
 Class: shotgun.
 Location/Qualifiers
 1..592
 /organism="Photorhabdus luminescens"
 /strain="W14"
 /db_xref="taxon:29488"
 /clone="pLG01581"
 /clone_lib="Photorhabdus luminescens strain W14 M13
 library"
 /dev_stage="primary phase variant"
 /note="Genomic DNA from strain W14 was size selected (1-2 kb)
 and then cloned into M13 Janus."
 146 a 140 c 112 g 191 t 3 others

Query Match	9.9%	Score 59.4	DB 13	Length 592
Best Local Similarity	51.3%	Pred. No. 146	0	
Matches 138	Conservative 0	Mismatches 131	Indels 0	Gaps 0
QY	210	tgacctggtttcgaagatcgccagaagtcacactgacaaagctcagggccgagctgc	263	
Db	293	TGAATTGATTGATTCATTCGACAGAAAGCAACGACCTCGATTAAGCTCAGCGCAAAGGCTGGC	234	
QY	270	tgttaacgacctccagatgtgtctctcgaggtctatgaagtcgcggcgaagggctgaagct	329	
Db	223	TCTGGATTCAACTTTGATGTGATTCATACCGAATCCCTGGAAGGGGGGATTCAGTCAACT	174	
QY	330	caaccgaccttcttcgcgtcgtcagcgctcaagcgcccgctcgacccggccgcacaccgcg	389	
Db	173	GGTAGGTTTTGGGACCTTTCAAAGTTAAACATCTCTGAGAACTACGCGGTCTGTAACCCGCA	114	
QY	390	cactgcgcgacgaattacatctccggtcttcactcagtcgcttcatccgctcgtctccct	449	
Db	113	AACCGGTAAAGAATATTCAATTCGCGGGCTGGAACAGTACACTGGTTTCTCTGCTGTAAGGC	54	
QY	450	gctcgaagaagggcggtcacgcagctgacctt	478	
Db	53	TTTGAAAGACGCAAGTTAAATATTTCAATT	25	

FEATURES	source
BASE COUNT	64 a 128 c 116 g 68 t
ORIGIN	
Query Match	9.2%; Score 55; DB 10; Length 376;
Best Local Similarity	52.9%; Pred. No. 2,3e-05;
Matches 118; Conservative 0; Mismatches 105; Indels 0; Gaps 0;	
Db	
204	caagctgacctcgcttcgaagatgccacagaagtcacaactctgaccaaagctcaaggccga 263
151	CAAGGCTTAACTTGTGGAGGGCGATGCTACATGAGGTGGAGTGAACCAAGAGCTGCTGC 210
Qy	
264	gacgcgtgtgaagcgcctccagatgctgttcgtcgaagctatgaagctccgcggaagcct 323
Db	
211	GAAAGCCTTATACATCCCGATGATGGCGGCAATGAGACCCCTCATCAACAGGTGACCGCT 270
Qy	
324	gaagctcaacgcgcctgtctctcgcgtcgaagcgcgtcaagcgcgcgcgtctgcacacgcgcga 383
Db	
271	GACCATCGTCGGGCTTCCGCGACGTTTGAAGTGTCGCGAGCGGCGCTGCGCCGACGGCGCAA 330
Qy	
384	ccgcgcgaactggcgagcagatgtgacattccgcttcctcaagc 426
Db	
331	CCCCAGCACCGCGCGGTCTCCACAGTCCGCCGACGACGAAGGCC 373
RESULT	
9	
LOCUS	AO989976 627 bp DNA GSS 14-AUG-2000
DEFINITION	Rec00661 Photorhabdus luminescens strain W14 M13 library
ACCESSION	AO989976
VERSION	AO989976.1 GI:9648570
KEYWORDS	GSS.
SOURCE	Photorhabdus luminescens.
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
1 (bases 1 to 627)	french-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T., Daborn, P.J., Bowen, D., and Blattner, F.R.	A genomic sample sequence of the entomopathogenic bacterium <i>Photobacterium luminescens</i> W14: potential implications for virulence	Appl. Environ. Microbiol.	66 (8), 3310-3329 (2000)	20378633
					Contact: french-Constant RH Department of Biology and Biochemistry University of Bath South Building, Bath BA2 7AY, UK Tel: (44) 1225 826621 Fax: (44) 1225 826779 Email: bsr@icbath.ac.uk This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli M13 genome) please see french-Constant et al. 2000, Nucleic Acids Res. Seq primer: M13 Forward Class: shotgun.
FEATURES	SOURCE	location/Qualifiers			
		1..627			
		/organism="Photobacterium luminescens"			
		/strain="W14"			
		/db_xref="taxon:29488"			
		/clone="PIC00661"			
		/clone.lib="Photobacterium luminescens strain W14 M13 library"			
		/dev_stage="primary phase variant"			
		/note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."			
BASE COUNT	189 a	131 c	137 g	166 t	4 others
ORIGIN					
Query Match	8.6%	Score 51.6	DB 13	Length 627	
Best Local Similarity	50.8%	Pred. No. 0.00027			
Matches 123	Conservative 0	Mismatches 119	Indels 0	Gaps 0	
OY	202	aaccagctgcacctgcttcggaagtcgcccgaagtcacaactgacaaagctcagcc	261		
DB	229	ATAAGACTGATTTAGTTGATGATCAATCCGACAAACACAGACTTAAAGCTAAGCA	288		
OY	262	gagctctgttaacgacctccagatgtgttcgtcgaagctatgaagtcgcgcaagcc	321		
DB	289	AAGCTGCTCTGATCAATCACTTTGATGATCAATCACCGAATCCCTGAAGCGGATTTCA	348		
OY	322	ctgaagctcaacgacctgttcctcgtctgagcgctcaagcgcccgctcgcacgcgcgc	381		
DB	349	GTACAACTGGTAGGTTTGGGACTTTAAAGTTAAACCATCGTGCAGAAACGTCGGTCT	408		
OY	382	aaccgcgcacatggcgagcagatgacatccgcgcttcctacgcgcgttcgataccgcg	441		
DB	409	AACCGCAACCGGAAAGAAATTAACATCCGCGGTGCAAAAGTACCTGTTCTCTCT	468		
OY	442	gg 443			
DB	469	GG 470			
RESULT 10					
LOCUS	A2934085	800 bp	DNA	GSS	24-Apr-2001
DEFINITION	Bt_Ba000101Db B. japonicum BAC library				Bradyrhizobium japonicum genomic, DNA sequence.
ACCESSION	A2934085				
VERSION	A2934085.1	GI:13776145			
KEYWORDS	GSS.				
SOURCE	Bradyrhizobium japonicum.				
ORGANISM	Bradyrhizobium japonicum				
	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;				
	Bradyrhizobium group; Bradyrhizobium.				

Query Match	8.3%	Score 49.8	DB 13	Length 800
Best Local Similarity	54.8%	Pred. No. 0.00097		
Matches 121	Conservative	0	Mismatches 97	Indels 3
			Gaps	1
184	ggaatgcttattatgcatatacaacaagctgacacctgcgtttctcgaagaatgcc---	cagaagtcc	240	
Db				
388	ggggcctatctctctgcattcatcaaacccaactgtttcagcctatgcgcagacacacccg	447		
Qy				
241	aacttgaccgaagagctcgaagccgctgctgttaacgctctccagaatgctgtcgag	300		
Db				
448	catctgtracacgacgggaatgtcgagaaacatttgtaattgccaatctcgamaagatcgtagcg	507		
Qy				
301	gcatatgaagtcgagcgaagccttgaagctcaacggcctgtttccgtctgagcgctcaag	360		
Db				
508	gctcttcgcacacgcggggaatgcgctcagatgtggcgcccttcgtcggtcgaagcat	567		
Qy				
361	cgcccgagctgcacgcgcgcgaaccgcgcgaactgtgcagca	401		
Db				
568	cgccctgcacacgcgcgcgagctgaattccgcgcacccgcgcga	608		
RESULT 11				
LOCUS	BF483682			
DEFINITION	WHE2336.F03.L0663 wheat pre-anthesis spike cDNA library Triticum aestivum cDNA clone WHE2336.F03.L06, mRNA sequence.			
ACCESSION	BF483682			
VERSION	BF483682.1			
KEYWORDS	EST.			
SOURCE	wheat.			
ORGANISM	Triticum aestivum			
REFERENCE	Eukaryotes: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae; Pooidaeae; Triticeae; Triticum.			
AUTHORS	1 (bases 1 to 663)			
TITLE	Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han,P.S., Hsia,C.C., Kang,Y., Iazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and Tong,J.C.			
JOURNAL	The structure and function of the expressed portion of the wheat genomes - Pre-anthesis spike cDNA library			
COMMENT	Unpublished (2000)			
	Contact: Olin Anderson			
	US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center			

Query Match	7.7%	Score 46.2	DB 11	Length 663
Best Local Similarity	50.7%	Pred. No. 0.01	Mismatches 108	Indels 0
Matches 111	Conservative	0	Mismatches 108	Indels 0
				Gaps 0
OY	226	atgccacgaagatccaactgacacgtaccgaagctcaagccgagcgctgtgtttaacgcttcag	285	
Db	92	ATCACACACATATGGCAATGACCCACCAAGATATGGCGCGCAAAATGTGTCGGCGCCCGCCGCA	151	
OY	286	gatgtgttcctgcgagggcgtatgaagtcgsgaagggccgaagctccacccgctgtttcc	345	
Db	152	GAACCCGTTACCAAGACTGCAAGCTGCCAGCCCGGCTTCMAAGGCCGCGACAAATCCGT	211	
OY	346	gtcagcgcgctcaaacgcgccggtctgcacacgcgcgaaccccgcgactgcgagacagatt	405	
Db	212	GCCGAGCCCGCTGCGCGCGCCAGTGCAGCCGAGCCGCGCAATATCGCTGCGAGCCCGCT	271	
OY	406	gaacttcgagcttctctaagcggtctgtctatctccgctgc	444	
Db	272	GCGTGGCCAGCTGCCGAAGCCGCGCAAAATCCGCTGCC	310	
RESULT 12				
LOCUS	BG309663			
DEFINITION	BG309663	831 bp	mRNA	EST
				22-FEB-2001
				HYSMC0011M08f Hordeum vulgare seedling shoot EST library
				HYSDNA0003 (Etiolated and unstressed) Hordeum vulgare cDNA clone
				HYSMC0011M08f, mRNA sequence.
ACCESSION	BG309663			
VERSION	BG309663.1	GI:13110510		
KEYWORDS	EST.			
SOURCE	barley.			
ORGANISM	Hordeum vulgare			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;			
	Triticeae; Hordeum.			
	1 (bases 1 to 831)			
REFERENCE	Wing,R., Close,T.J., Kleinbols,A., Wise,R., Begum,D., Fritsch,D., Yu			
AUTHORS	,Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo			
	,T., Saski,C., Schwartzbeck,J., Simons,J., Choi,D.W., Main,D. and			
	Wood,T.			
TITLE	Development of a genetically and physically anchored EST resource			
	for barley genomics			

Email: mmp@attetuga.edu
Sequences have been trimmed to exclude PolyA, vector and regions below phred quality 16. The threshold for highest quality sequence is 20.

Seq primer: PolyTMix
High quality sequence stop: 488
POLYA-No.

FEATURES
source location/Qualifiers

1..493
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_11b="Dark Grown 1 (DG1)"
/note="Organ: 5-day-old dark-grown seedlings; Vector: lambda zap; Site_1: XhoI; Site_2: EcoRI. The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."
BASE COUNT 78 a 152 c 160 g 103 t
ORIGIN

Query Match 7.18; Score 42.6; DB 10; Length 493;
Best Local Similarity 47.9%; Pred. No. 0.1;
Matches 123; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 194 tggcattacaagaagctcgaactcgttcgaagaatcgcccaagaatccaactgaaccagg 253
DB 118 TGGGTTTCGACCCGAGAGCCGCGGTGTTCACGCGGAGAGATGACGCGCATGAGACGC 177
QY 254 ctccgagcgaagctcgtcttaacgccttcacagatgtctcgaagctatgaatccg 313
DB 178 ACTGGGAAGGTCAGCCGCCGCCGCCGAGAGAGAGAGATGAGAGAGAGAGTTCA 237
QY 314 gcgaaagcctgaagctacacgcctgttcctcgcgtgagcgcgtaagcgccgcgtcga 373
DB 238 TGCAGGCTTCACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 297
QY 374 ccggccgcaacccgc 433
DB 298 CTTGGTAAACAAATCTCAAGTCGCGCAGCAAGTTGTCAGTCCGCCGCCGCGTGGTAATC 357
QY 434 tctcgcgtgcctccctg 450
DB 358 TAAACCCCGACCGCCCG 374

RESULT 15
BI099174 578 bp mRNA EST 26-JUN-2001
LOCUS IP1.38.G12.g1_A002 Immature pannicle 1 (IP1) Sorghum bicolor cDNA,
DEFINITION mRNA sequence.
ACCESSION BI099174
VERSION BI099174.1 GI:14570756
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 578)
Klein,R.R., Cordonnier-Pratt,M.-M., Gingle,A., Sudman,M. and Pratt
L.H.
An EST database from Sorghum: developing preanthesis pannicles
Unpublished (2001)
Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmp@attetuga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below phred quality 16. The threshold for highest quality sequence is 20. Three-prime sequences, which are obtained with PolyTMix or T7

sequencing primer, are presented as the reverse complement.

Seq primer: T7
High quality sequence start: 2
High quality sequence stop: 578
POLYA-No.

FEATURES
source location/Qualifiers

1..578
/organism="Sorghum bicolor"
/cultivar="Brx623"
/db_xref="taxon:4558"
/clone_11b="Immature pannicle 1 (IP1)"
/note="Organ: Developing preanthesis pannicles; Vector: Bluescript II SK(-) from lambda zap II; Site_1: XhoI; Site_2: EcoRI. The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."
BASE COUNT 102 a 164 c 172 g 140 t
ORIGIN

Query Match 7.18; Score 42.6; DB 11; Length 578;
Best Local Similarity 47.9%; Pred. No. 0.11;
Matches 123; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 194 tggcattacaagaagctcgaactcgttcgaagaatcgcccaagaatccaactgaaccagg 253
DB 109 TGGGTTTCGACCCGAGAGCCGCGGTGTTCACGCGGAGAGATGACGCGCATGAGACGC 168
QY 254 ctccgagcgaagctcgtcttaacgccttcacagatgtctcgaagctatgaatccg 313
DB 169 ACTGGGAAGGTCAGCCGCCGCCGCCGAGAGAGAGAGATGAGAGAGAGAGTTCA 228
QY 314 gcgaaagcctgaagctacacgcctgttcctcgcgtgagcgcgtaagcgccgcgtcga 373
DB 229 TGCAGGCTTCACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 288
QY 374 ccggccgcaacccgc 433
DB 289 CTTGGTAAACAAATCTCAAGTCGCGCAGCAAGTTGTCAGTCCGCCGCCGCGTGGTAATC 348
QY 434 tctcgcgtgcctccctg 450
DB 349 TAAACCCCGACCGCCCG 365

Search completed: December 20, 2001, 19:20:31
Job time: 6407 sec

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2002, 11:24:49 ; Search time 43.6636 Seconds
(without alignments)
1348.538 Million cell updates/sec

Title: US-09-816-391a-1_COPY_1_192

Sequence: 1 gctggcgcgcgccatga.....tgaccacaagaagatgcttt 192

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/5B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29.2	15.2	3092	1	US-08-426-627-3
2	29.2	15.2	3126	2	US-08-477-396A-3
3	29.2	15.2	3253	1	US-08-426-627-5
4	27.8	14.5	730	4	US-08-943-731-128
5	27.8	14.5	1257	1	US-08-330-154-1
6	27.8	14.5	1269	4	US-09-134-001C-2393
7	27.8	14.5	24183	4	US-08-943-731-3
8	27.4	14.3	7650	4	US-09-221-017B-911
9	27.2	14.2	168575	4	US-09-426-290-1
10	27.2	14.1	1009	6	5223394-8
11	27.2	14.1	1244	2	US-08-204-288-3
12	27.2	14.1	3182	1	US-08-188-582-12
13	27.2	14.1	3182	1	US-08-646-715-11
14	27.2	14.1	3852	1	US-08-306-546C-1
15	27.2	14.1	3852	2	US-08-530-524A-1
16	26.8	14.0	875	4	US-09-221-017B-671
17	26.8	14.0	1710	4	US-08-630-915A-189
18	26.8	14.0	42571	4	US-09-810-347-3
19	26.6	13.9	3484	4	US-09-308-090-1
20	26.2	13.6	1610	4	US-09-347-803-19
21	26.2	13.6	2005	4	US-08-936-165A-102
22	26.2	13.6	2559	4	US-09-071-035-325
23	26.2	13.6	2559	4	US-09-071-035-329
24	26.2	13.6	2559	4	US-09-071-035-333
25	26.2	13.6	2801	3	US-08-747-221B-30
26	26.2	13.6	2801	3	US-08-747-221B-32
27	26.2	13.6	2801	4	US-09-005-051-30

28	26.2	13.6	2801	4	US-09-005-051-32	Sequence 32, Appl
29	26.2	13.6	2836	3	US-08-747-221B-24	Sequence 24, Appl
30	26.2	13.6	2836	3	US-08-747-221B-26	Sequence 26, Appl
31	26.2	13.6	2836	4	US-09-005-051-24	Sequence 24, Appl
32	26.2	13.6	2836	4	US-09-005-051-26	Sequence 26, Appl
33	26.2	13.6	5506	4	US-09-004-838-93	Sequence 93, Appl
34	26.2	13.5	490	4	US-09-222-575-95	Sequence 95, Appl
35	26.2	13.5	1851	4	US-09-134-001C-1781	Sequence 1781, Ap
36	25.8	13.4	885	2	US-08-222-719-7	Sequence 7, Appl
37	25.8	13.4	885	2	US-08-470-925-7	Sequence 7, Appl
38	25.8	13.4	885	2	US-08-471-613-7	Sequence 7, Appl
39	25.8	13.4	885	5	PCT-US93-10443-7	Sequence 7, Appl
40	25.8	13.4	2144	4	US-09-221-017B-731	Sequence 731, App
41	25.8	13.4	18436	4	US-08-961-827-87	Sequence 87, Appl
42	25.6	13.3	735	3	US-09-193-191-3	Sequence 3, Appl
43	25.6	13.3	1062	4	US-09-371-913A-5	Sequence 5, Appl
44	25.6	13.3	2172	1	US-08-158-232-50	Sequence 50, Appl
45	25.6	13.3	2172	1	US-08-611-928-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1
US-08-426-627-3
Sequence 3, Application US/08426627
Patent No. 5756664
GENERAL INFORMATION:
APPLICANT: Amann, Egon
APPLICANT: Otawara-Hamamoto, Yoko
APPLICANT: Kikuno, Reiko
APPLICANT: Takeshita, Sunao
APPLICANT: Tezuka, Kenichi
TITLE OF INVENTION: No. 5756664 Protein with Bone Formation
TITLE OF INVENTION: Ability and Process for its Production.
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,627
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/036,841
FILING DATE: 25-MAR-1993
APPLICATION NUMBER: JP 4-71501
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hammond, Alan W.
REGISTRATION NUMBER: 35,178
REFERENCE/DOCKET NUMBER: 02481-1285-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3092 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:

ORGANISM: Homo sapiens
TISSUE TYPE: Placenta
FEATURE:
NAME/KEY: CDS
LOCATION: join(38..2375)
FEATURE:
NAME/KEY: mat.peptide
LOCATION: join(101..2375)
US-08-426-627-3

Query Match 15.2%; Score 29.2; DB 1; Length 3092;
Best Local Similarity 59.8%; Pred. No. 1.4;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 89 AACCTTATTAACGGGGTTTTCGCAAGAACATGCGCTAGTATCATGTATGACAAACATG 148
DB 626 AACCATATTCCTAATGGGGTGTCACTGTTAATGTGTCGATCATATCCATGGGAACACAG 685
QY 149 GACTAGCAAAAGTGTGTCC 170
DB 686 ATTGCAACAATGGTGTGTCC 707

RESULT 2

US-08-477-396A-3
Sequence 3, Application US/08477396A
Patent No. 5872235
GENERAL INFORMATION:
APPLICANT: Chen, Ian Bo
APPLICANT: Bao, Shideng
APPLICANT: Liu, Yuan
TITLE OF INVENTION: A NOVEL TUMOR MARKER AND NOVEL METHOD OF
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,396A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US-08/146,488
FILING DATE: 29-OCT-1993
APPLICATION NUMBER: US 08/448,388
FILING DATE: 28-MAY-1996
PRIOR APPLICATION DATA: PCT/US94/12502
APPLICATION NUMBER: 31-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Heine, Holliday C.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCL-333BX
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO

ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 43..2376
US-08-477-396A-3

Query Match 15.2%; Score 29.2; DB 2; Length 3126;
Best Local Similarity 59.8%; Pred. No. 1.4;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 89 AACCTTATTAACGGGGTTTTCGCAAGAACATGCGCTAGTATCATGTATGACAAACATG 148
DB 649 AACCATATTCCTAATGGGGTGTCACTGTTAATGTGTCGATCATATCCATGGGAACACAG 708
QY 149 GACTAGCAAAAGTGTGTCC 170
DB 709 ATTGCAACAATGGTGTGTCC 730

RESULT 3

US-08-426-627-5
Sequence 5, Application US/08426627
Patent No. 5756664
GENERAL INFORMATION:
APPLICANT: Amann, Egon
APPLICANT: Otawara-Hamamoto, Yoko
APPLICANT: Kikuno, Reiko
APPLICANT: Takeshita, Sunao
APPLICANT: Tezuka, Kenichi
TITLE OF INVENTION: No. 5756664e1 Protein with Bone Formation
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,627
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/036,841
FILING DATE: 25-MAR-1993
APPLICATION NUMBER: JP 4-71501
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hammond, Alan W.
REGISTRATION NUMBER: 35,178
REFERENCE/DOCKET NUMBER: 02481-1285-00000
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3253 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: osteosarcoma
FEATURE:
NAME/KEY: CDS

Query Match	14.58;	Score 27.8;	DB 1;	Length 1257;
Best Local Similarity	62.08;	Pred. No. 3;		
Matches 44; Conservative	0;	Mismatches 27;	Indels 0;	Gaps 0;

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Monroy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 911:
SEQUENCE CHARACTERISTICS:
LENGTH: 7650 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...7650
US-09-221-017B-911

Query Match 14.3%; Score 27.4; DB 4; Length 7650;
Best Local Similarity 53.2%; Pred. No. 8.4;
Matches 58; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 39 AACTTCGATTCGCTATTTTCATACCTTGGGGAATTAATGTGAACCCCTTATA 98
DB 1504 AACTGCCCAAACTTCGATTCGCTACCTCCGCTTTTAAACAGGATCGGCCCTTAA 1445
QY 99 AAACGGGGTTTTCGCAAAACATGCGCTAGTATCATTTGATGACAACAT 147
DB 1444 AAGCGGCTCTCTTCGCGCAAAAGAGCTGAGCATCTCGATGCGACCTT 1436

RESULT 9
US-09-426-290-1
Sequence 1, Application US/09426290
Patent No. 6410712
GENERAL INFORMATION:
APPLICANT: Berigild Ran Olafsdottir
APPLICANT: Jeffrey Gulcher
TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
FILE REFERENCE: 2345.2001-000
CURRENT APPLICATION NUMBER: US/09/426,290
CURRENT FILING DATE: 1999-10-25
NUMBER OF SEQ ID NOS: 24
SOFTWARE: FastSeq for Windows version 4.0
SEQ ID NO 1
LENGTH: 168575
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (21181)...(21403)
NAME/KEY: CDS
LOCATION: (95252)...(95430)
NAME/KEY: CDS
LOCATION: (101753)...(101966)
NAME/KEY: CDS
LOCATION: (110324)...(110439)
NAME/KEY: CDS
LOCATION: (124058)...(124278)
NAME/KEY: CDS
LOCATION: (127009)...(127130)

NAME/KEY: CDS
LOCATION: (128910)...(129139)
US-09-426-290-1

Query Match 14.2%; Score 27.2; DB 4; Length 168575;
Best Local Similarity 53.8%; Pred. No. 34;
Matches 56; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 56 TATTTCAATACCTTGGGGAATATAGTGAACCCCTTATAAAGCGGCTTCGCA 115
DB 109602 TATTCATATGATGATGAGAAATTAATGTATTAAGCCCTTAATTCATTTTATTA 109661
QY 116 GAAACATGCGCTAGTATTCATTCATGACAACATGACTAAGCAAA 159
DB 109662 AAGTATATATACATGATTTTATTAATAAGCATACACACACACCA 109705

RESULT 10
5223394-8
Patent No. 5223394
TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
LINKAGE SIGNAL SEQUENCE
NUMBER OF SEQUENCES: 12
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/335,688
FILING DATE: 10-APR-1989
SEQ ID NO: 8
LENGTH: 1009
5223394-8

Query Match 14.1%; Score 27; DB 6; Length 1009;
Best Local Similarity 57.8%; Pred. No. 5.2;
Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 74 GGAATATAGATGTAATAACCTTATATAAGCGGGTTTGCAGAAACATGCGCTAGTANC 133
DB 777 GGAAGTGTCTGTAAAGCGGTTTCAAGACACAGATATGACATATACCATACATTAGC 836
QY 134 ATTGATGACAACATGAGCTAAGC 156
DB 837 AGTATTTACACATGATGTGTC 859

RESULT 11
US-08-204-288-3/c
Sequence 3, Application US/08204288
Patent No. 5959178
GENERAL INFORMATION:
APPLICANT: VAN DOORSSELAERE, Jan
APPLICANT: FRITIG, Bernhard J.M.
APPLICANT: INZE, Dirk G.
APPLICANT: JOUANIN, Lise
APPLICANT: KNIGHT, Mary E.
APPLICANT: VAN MONTAGU, Marc
APPLICANT: LEGRAND, Michel
TITLE OF INVENTION: MODIFICATION OF LIGNIN SYNTHESIS IN
PLANTS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D. C.
COUNTRY: U.S.A.
ZIP: 20005-3518
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/204,288
FILING DATE: 10-MAR-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9119279.9
FILING DATE: 10-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/01460
FILING DATE: 09-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, Paul N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 206860/SPE36543/USI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1244 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-204-288-3

Query Match 14.1%; Score 27; DB 2; Length 1244;
Best Local Similarity 62.7%; Pred. No. 5.6;
Matches 42; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

OY 57 ATTTCATACCTTGGCGGAATATGTGAAAACCTTTATAACGGCGGTTTCGCAG 116
DB 564 ATCTCAATACATGTGGTAATCAAGTTATACCTTAATAGAGGATGTTTGAGAGC 505
OY 117 AACATG 123
DB 504 AATCATG 498

RESULT 12
US-08-188-582-12/C
Sequence 12, Application US/08188582
Patent No. 5534410
GENERAL INFORMATION:
APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
APPLICANT: Dynlacht, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS.
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,582
FILING DATE: 28-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A

REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3182 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 972..3002
US-08-188-582-12

Query Match 14.1%; Score 27; DB 1; Length 3182;
Best Local Similarity 66.1%; Pred. No. 8.2;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

OY 2 CTGGCGCGGCGGCGCATGAGTGGCTGACAGCATATCTTGCTGATTCGTATTT 60
DB 1238 CTGGCGTGGAGCCATGAGTGGCTGACATCTTTAGCTTCAAGCGGTAGTCAATGT 1180

RESULT 13
US-08-646-715-12/C
Sequence 12, Application US/08646715
Patent No. 5637686
GENERAL INFORMATION:
APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
APPLICANT: Dynlacht, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS.
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,715
FILING DATE: 09-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/188,582
FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3182 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 972..3002
US-08-646-715-12

Query Match 14.1%; Score 27; DB 1; Length 3182;
Best Local Similarity 66.1%; Pred. No. 8.2;

Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 2 CTGGCGCGCGCGCATGAGTGGCTTGACAAACATATCTGCTGATTCGCTATTT 60
DB 1238 CTGGCGCGTGAAGCCATGAGTGGCTGACATCTTACTCTCAAGGCTAGCAATGT 1180

RESULT 14
US-08-306-546C-1
Sequence 1, Application US/08306546C
Patent No. 5605797

GENERAL INFORMATION:

APPLICANT: Friderich, Karen
APPLICANT: Jones, Margaret
APPLICANT: Chen, Hong
APPLICANT: Cavanagh, Kevin
TITLE OF INVENTION: Bovine Beta-Mannosidase Gene and Methods
TITLE OF INVENTION: of Use
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESSES:
ADDRESS: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: Michigan
COUNTRY: USA
ZIP: 48303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,546C
FILING DATE: September 15, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REGISTRATION NUMBER: 36,683
REFERENCE/DOCKET NUMBER: 6550-00003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810)641-1600
TELEFAX: (810)641-0270
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3852 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 75..2711
US-08-306-546C-1

Query Match 14.1%; Score 27; DB 1; Length 3852;
Best Local Similarity 56.0%; Pred. No. 8.8;

Matches 51; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 75 GAAATAGATGCGAAACCTTATAAAGCGGCTTTGGCAGAAACATCGCTAGTATCA 134
DB 3567 GATTGCGATGATGAGCCTGGACCAAGTGTATGCTTAACTTGAGATGAAACA 3626

QY 135 TTGATGACACATGAGCTAGCAAGAGTGT 165
DB 3627 TTTTAAAAACACTGATTAATTAATTGCT 3657

RESULT 15
US-08-530-524A-1
Sequence 1, Application US/08530524A
Patent No. 5837836

GENERAL INFORMATION:

APPLICANT: Friderich, Karen
APPLICANT: Jones, Margaret
APPLICANT: Chen, Hong
APPLICANT: Cavanagh, Kevin
TITLE OF INVENTION: Bovine Beta-Mannosidase Gene and Methods
TITLE OF INVENTION: of Use
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESSES:
ADDRESS: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: Michigan
COUNTRY: USA
ZIP: 48303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,524A
FILING DATE: September 19, 1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REGISTRATION NUMBER: 36,683
REFERENCE/DOCKET NUMBER: 6550-00003DVA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810)641-1600
TELEFAX: (810)641-0270
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3852 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 75..2711
US-08-530-524A-1

Query Match 14.1%; Score 27; DB 2; Length 3852;
Best Local Similarity 56.0%; Pred. No. 8.8;

Matches 51; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 75 GAAATAGATGCGAAACCTTATAAAGCGGCTTTGGCAGAAACATCGCTAGTATCA 134
DB 3567 GATTGCGATGATGAGCCTGGACCAAGTGTATGCTTAACTTGAGATGAAACA 3626

QY 135 TTGATGACACATGAGCTAGCAAGAGTGT 165
DB 3627 TTTTAAAAACACTGATTAATTAATTGCT 3657

Search completed: December 25, 2002, 12:58:44
Job time : 144.664 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2002, 11:24:48 ; Search time 46.0561 Seconds
(without alignments)
1693.523 Million cell updates/sec

Title: US-09-816-391A-1_COPY_1_192

Perfect score: 192
Sequence: 1 gctggcgccgagcgagcatga.....tgaccacgaagatgcttt 192

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 363474 seqs, 203117208 residues

Total number of hits satisfying chosen parameters: 726948

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	192	100.0	600	10	US-09-816-391A-1
2	30.2	15.7	515	10	US-09-864-761-16206
3	29.2	15.2	3202	10	US-09-925-301-471
4	29.2	15.2	6465	10	US-09-954-456-193
5	28.8	15.0	413	10	US-09-867-701-2315
6	28.8	15.0	419	10	US-09-783-590-663
7	28.4	14.8	345	10	US-09-974-300-3770
8	28.4	14.8	75899	10	US-09-854-883-243
9	27.6	14.4	727	10	US-09-770-149-122
10	27.4	14.3	484	10	US-09-747-155-261
11	27.4	14.3	2000	9	US-09-938-842A-448
12	27.4	14.3	202001	10	US-09-734-674-2306
13	27	14.1	389	10	US-09-867-701-2306
14	27	14.1	1577	10	US-09-864-761-30755
15	27	14.1	1870	9	US-09-938-842A-3414
16	27	14.1	1955	10	US-09-864-761-14202
17	27	14.1	2725	10	US-09-864-824A-243
18	27	14.1	2725	10	US-09-880-107-2181
19	27	14.1	14070	9	US-09-870-759-127

20	27	14.1	14635	10	US-09-880-107-2287	Sequence 2287, Ap
21	26.8	14.0	1710	10	US-09-879-957-189	Sequence 189, Ap
22	26.8	14.0	1782	10	US-09-974-300-2801	Sequence 2801, Ap
23	26.6	13.9	182	10	US-09-783-590-4530	Sequence 4530, Ap
24	26.4	13.7	2324	10	US-09-771-161A-74	Sequence 74, Ap
25	26.4	13.7	2544	10	US-09-771-161A-74	Sequence 74, Ap
26	26.4	13.7	32038	10	US-09-764-878-292	Sequence 723, Ap
27	26.2	13.6	176	10	US-09-867-701-8723	Sequence 8723, Ap
28	26.2	13.6	855	10	US-09-974-300-998	Sequence 998, Ap
29	26.2	13.6	2000	9	US-09-938-842A-22759	Sequence 2759, Ap
30	26.2	13.6	2005	10	US-09-939-980-102	Sequence 102, Ap
31	26.2	13.6	4527	10	US-09-901-940-3	Sequence 102, Ap
32	26.2	13.6	4594	10	US-09-070-922A-230	Sequence 230, Ap
33	26.2	13.6	63000	10	US-09-780-172-18	Sequence 18, Ap
34	26.2	13.5	402	9	US-09-933-797-133	Sequence 133, Ap
35	26	13.5	490	10	US-09-604-287A-95	Sequence 95, Ap
36	26	13.5	490	10	US-09-339-338-95	Sequence 95, Ap
37	26	13.5	490	12	US-10-007-803-95	Sequence 95, Ap
38	26	13.5	546	10	US-09-777-564-1613	Sequence 1613, Ap
39	26	13.5	814	9	US-10-001-857-5	Sequence 5, Ap
40	26	13.5	900	10	US-09-974-300-2838	Sequence 2838, Ap
41	26	13.5	1365	12	US-10-073-256-49	Sequence 49, Ap
42	26	13.5	302250	10	US-09-962-832-154	Sequence 154, Ap
43	25.8	13.4	201	10	US-09-783-590-5925	Sequence 5925, Ap
44	25.8	13.4	391	10	US-09-864-761-4151	Sequence 4151, Ap
45	25.8	13.4	397	10	US-09-960-352-2723	Sequence 2723, Ap

ALIGNMENTS

RESULT 1
US-09-816-391A-1
Sequence 1, Application US/09816391A
Patent No. US20020054865A1
GENERAL INFORMATION:
APPLICANT: FUJIMORI, Minoru
APPLICANT: TANIGUCHI, Shunichiro
APPLICANT: AMANO, Jun
APPLICANT: YAZAWA, Kazuyuki
APPLICANT: KANO, Yasunobu
APPLICANT: NAKAMURA, Toshitsuyuki
APPLICANT: SAKAKI, Takayuki
TITLE OF INVENTION: Anaerobic bacterium as a drug for cancer gene therapy
FILE REFERENCE: 2001-WM/01736
CURRENT FILING DATE: US/09/816,391A
CURRENT FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: JP 00/287688
PRIOR FILING DATE: 2000-09-21
NUMBER OF SEQ ID NOS: 3
SEQ ID NO 1
LENGTH: 600
TYPE: DNA
ORGANISM: Bifidobacterium longum
NAME/KEY: CDS
LOCATION: (193)..(471)
US-09-816-391A-1

Query Match 100.0% Score 192, DB 10; Length 600;
Best Local Similarity 100.0%; Pred. No. 3, 1e-57;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCTGGCGCGCGCCATGAGTGGCTTGACACATTAATCTGTGATGCTGCTATTT 60
DB 1 GCTGGCGCGCGCGCCATGAGTGGCTTGACACATTAATCTGTGATGCTGCTATTT 60
QY 61 TCAATACCTTCGGGGAATAGATGTGAATCCCTTATTAACCGGGGTTTTCGAGAAC 120
DB 61 TCAATACCTTCGGGGAATAGATGTGAATCCCTTATTAACCGGGGTTTTCGAGAAC 120
QY 121 ATGGCTAGATCATTTATGACACATGACACTAAGCAAAAGTGGTTCCTCCGACCCAA 180
DB 121 ATGGCTAGATCATTTATGACACATGACACTAAGCAAAAGTGGTTCCTCCGACCCAA 180

```

Db      121 ATGCAGTAGATCATTCATGATGACAAACATGGACTAGCAAAAAGTCTGTGCCCTGACCA 180
QY      181 GAAGGATGCTTT 192
          |||||
          |||||
Db      181 GAAGGATGCTTT 192

RESULT 2
US-09-864-761-16206
; Sequence 16206, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16206
; LENGTH: 515
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000062.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
US-09-864-761-16206

Query Match          15.7%; Score 30.2; DB 10; Length 515;
Best Local Similarity 49.7%; Pred. No. 0.56;
Matches 77; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

13 GCCATGATGGCTTGACAAACATTCCTGATGCTGATTCCTATTTTCAATACCTCG 72

```

```

Db      253  GCCCTTAAATGACCACTAGATATTTATCTTAGTGAAGTACAGTTTAAATCTTTA 312
OY      73  GGGAAATAGATGCAAAACCTTTATTAAGCGGGGTTTGGCAGAAACATGCGCTAGTAT 132
Db      313  TAGAAGTAGATGATMAACATTTCCACATTAATGTTTGTATGATTCATGCTGTGTTT 372
OY      133  CATGATGACACACATGAGCTAAGCAAAAGTCTTG 167
Db      373  GTTTACTGCTCTTTAAACTTTTAGAATCTTCG 407

RESULT 3
US-09-925-301-471
; Sequence 471, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 471
; LENGTH: 3202
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (3160)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (3180)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-471

Query Match      15.2%; Score 29.2; DB 10; Length 3202;
Best Local Similarity 59.8%; Pred. No. 3;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY      89  AACCTTTAAAGCGGGTTTGCAGAAACGCGCTAGTATCATTTGATGACACATG 148
Db      633  AACCATTTCTTAATGGGGTTGTCACTGTTAATGTGCTGAAATCATCATGGGAAACAG 682
OY      149  GACTAAGCAAAAGTCTGTGCC 170
Db      683  ATTGCAACAATGGGTGTGCC 704

RESULT 4
US-09-954-456-193
; Sequence 193, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923

```



```

; LOCATION: (315)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (332)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (335)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (338)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (340)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (348)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (357)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (375)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (380)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (381)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (390)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (397)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (413)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-663

Query Match          15.0%; Score 28.8; DB 10; Length 419;
Best Local Similarity 50.4%; Pred. No. 1.6;
Matches 60; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 42 TTGCTGATTCCTATTTTCAATACCTTCGGGAAATAGATGTGAAACCTTATTA 101
DB 175 TTTCCTTTGCTCTTAATCTCAACACTTTGGAGACCAAGGTGGGATGGCTGAAC 234

QY 102 CGCGGGTTTTCGAGAAACATGCGCTAGTATCATTTGAGACAACATGACTAAGCA 160
DB 235 CAGGAGTTTNTGANCACCCCTGGGAGNANCATTAAGTGAGNACCCTTTTCAANAAA 293

RESULT 7
US-09-974-300-3770/c
; Sequence 3770, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Id Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085, 500-US
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US/09/974,300
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481

```

```

; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 3770
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
; US-09-974-300-3770

Query Match          14.8%; Score 28.4; DB 10; Length 345;
Best Local Similarity 49.3%; Pred. No. 2;
Matches 74; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 17 ATGAAGTGGCTTGAACAACATATCTGTGATTCGTCATTTTCAATACCTTCGGGGA 76
DB 258 ATGAATTTTATGGACGATGATCTTATCCGTCTTTTGTCTATTTCCGCA 199

QY 77 AATAGATGTGAACCCCTTATPAACGCGGTTTTCGAGAACATGCGTATATCAT 136
DB 198 AAGAGAAATTTTACAAATTTCAACAAGTTGTTCAAAATTTCAATATGAAACGCTTT 139

QY 137 GATGACATGAGCTAGCAAAAGTCTT 166
DB 138 TGCTGTAAATAGATTAAGAACGCTTTT 109

RESULT 8
US-09-854-883-243/c
; Sequence 243, Application US/09854883
; Patent No. US20020053479A1
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; APPLICANT: Jacqueline Wyatt
; APPLICANT: Susan M. Freier
; APPLICANT: Brett P. Monia
; APPLICANT: Madeline M. Butler
; APPLICANT: Robert McKay
; TITLE OF INVENTION: ANTISENSE MODULATION OF PTP1B EXPRESSION
; FILE REFERENCE: ISPH-0576
; CURRENT APPLICATION NUMBER: US/09/854,883
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 09/629,644
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 09/487,368
; PRIOR FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 389
; SEQ ID NO 243
; LENGTH: 75899
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; US-09-854-883-243

Query Match          14.8%; Score 28.4; DB 10; Length 75899;
Best Local Similarity 58.1%; Pred. No. 26;
Matches 50; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 57 ATTTCATACCTTCGGGGAATAGATGAAACCTTATTAACGGGGTTTCGAG 116
DB 16542 ATTTTCACTTCCTCCGAAATCAGAAAAAAGCATATATCAACTGGGCTTCTGCA 16483

QY 117 AAACATGCGCTAGTATCATTTGATGAC 142
DB 16482 GAGTTTGGCTTCTATCCAGATGCC 16457

RESULT 9
US-09-770-149-122/c
; Sequence 122, Application US/09770149
; Patent No. US20020059663A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.

```

APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Mathew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Woessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Krickler, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hubban, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2024 (PARA-013PRV)
CURRENT APPLICATION NUMBER: US/09/770,149
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/178,506
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 122
LENGTH: 727
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-149-122

Query Match 14.4%; Score 27.6; DB 10; Length 727;
Best Local Similarity 63.6%; Pred. No. 5.3;
Matches 42; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 29 GACAGCATATCTGTCGTGCTATTTCAATACCTCGGGGAATAGATGTGAA 88
DB 350 GACACATATCTGTCGTGCTGCTGACTCAAGACAGTAAAGATATATGCA 291

QY 89 AACCT 94
DB 290 ATACTT 285

RESULT 10
US-09-747-155-261/c
Sequence 261, Application US/09747155
Patent No. US20020151692A1
GENERAL INFORMATION:
APPLICANT: Rouquier, Sylvie
APPLICANT: Giorgi, Dominique
TITLE OF INVENTION: No. US20020151692A1 Polypeptides and Nucleic Acids Encoding Same
FILE REFERENCE: 19904-008 (C009B634US)
CURRENT APPLICATION NUMBER: US/09/747,155
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/171,746
PRIOR FILING DATE: 1999-12-22
NUMBER OF SEQ ID NOS: 431
SOFTWARE: PatentIn version 3.0
SEQ ID NO 261
LENGTH: 484
TYPE: DNA
ORGANISM: Eulemur rubriventer
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(484)
OTHER INFORMATION: Taxon = 34829; gene = EKU159; pseudogene; Accession DDBJ/EMBL/Gen
US-09-747-155-261

Query Match 14.3%; Score 27.4; DB 10; Length 484;
Best Local Similarity 51.2%; Pred. No. 5.1;
Matches 64; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 19 GAAGTGCTGACAGCATATCTGTCGTGCTATTTCAATACCTCGGGGAAA 78

DB 263 GCAGAGATTCACCAATATTAACATACCTTGCACCTGCCATGATATAGCATTCATCGGAAA 204
QY 79 TAGATGTGAAAACCTTTATAAACCAGGGTTTCCAGAAACATCGCTAGTATGTA 138
DB 203 TATGTGAGCATATGAGTTAAGAGCCAGAGATCTACAGGTATATGATCAATGATATAT 144
QY 139 TGACA 143
DB 143 TGCCA 139

RESULT 11
US-09-938-842A-4498/c
Sequence 4498, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Krepes, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 4498
LENGTH: 2000
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-4498

Query Match 14.3%; Score 27.4; DB 9; Length 2000;
Best Local Similarity 65.6%; Pred. No. 10;
Matches 40; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 31 CAAGCATATCTGTCGTGCTATTTCAATACCTCGGGGAATAGATGTGAAA 90
DB 1332 CAAGCATATCTGTCGTGCTATTTCAATATTAATATGACATATATATACAAA 1473

QY 91 C 91
DB 1472 C 1472

RESULT 12
US-09-734-674-3/c
Sequence 3, Application US/09734674
Patent No. US20020081648A1
GENERAL INFORMATION:
APPLICANT: Wei, Ming-Hui et al
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
FILE REFERENCE: CL001018
CURRENT APPLICATION NUMBER: US/09/734,674
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 202001
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(202001)

```
OTHER INFORMATION: n = A,T,C or G
US-09-734-674-3

Query Match      14.3%; Score 27.4; DB 10; Length 202001;
Best Local Similarity 62.3%; Pred. No. 91;
Matches 43; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 53 GCTATTTCATACCTTCGGGGAATAGATGCAAAACCTTATAAAGCGGCTTTC 112
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 44673 GACTATTATGAAACCTCTGTGCAAACAAATAGAAATCTAGAGAAATGATTAAGTTTC 44614
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 113 GCAGAAACA 121
    ||||| |||||
DB 44613 TCAGAAACA 44605

RESULT 13
US-09-867-701-2306/C
; Sequence 2306, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert.
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2306
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-867-701-2306

Query Match      14.1%; Score 27; DB 10; Length 389;
Best Local Similarity 60.0%; Pred. No. 6.4;
Matches 45; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY 35 CATAACTCTGATTCGCTATTTTCATACCTTCGGGGAATAGATGCAAAACCT 94
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 278 CACATGATGCTCATGCTGCTATTTTCATACCTTCGGGAGGCTTAGGTGGAGATCAC 219
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 95 TATTAACCGCGGTT 109
    ||||| |||||
DB 218 ATGACCCAGGGTT 204

RESULT 14
US-09-864-761-30755/C
; Sequence 30755, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30755
; LENGTH: 1577
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010872.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 9.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.2
; OTHER INFORMATION: NT HIT: M19828.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: R06764.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P04114, EVALUE 0.00e+00
US-09-864-761-30755

Query Match      14.1%; Score 27; DB 10; Length 1577;
Best Local Similarity 57.8%; Pred. No. 12;
Matches 48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

OY 85 TGAACCCCTTATATAAACCGCGGTTTCGCAGAACATGCGCTAGTATCATGTGACAA 144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 568 TAAAGATAGTTATGATTTACATGATTTGAAATAATGCTATTCCTAATATTTATGATGAAT 509
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 145 CATGACTAAGCAAAAGTCCTG 167
    ||||| ||||| ||||| |||||
DB 508 CATTGAAAATTAATAAGTCTTG 486

RESULT 15
US-09-938-842A-3414/C
; Sequence 3414, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCDP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
```

; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 3414
 ; LENGTH: 1870
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-3414

Query Match 14.1%; Score 27; DB 9; Length 1870;
 Best Local Similarity 53.3%; Pred. No. 14;
 Matches 57; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 QY 29 GACAAGCAATATCTGCTGATTCGTCATATTTCAATACCTTCGGGGAATAGATGTGAA 88
 DB 1599 GAAAAAAGTATCTTATATTTTAGGCCATATAATACCATATTTGAAATAGATATTTT 1540
 QY 89 AACCTTATATAACGGGGGTTTTCGCAGAAACATCGCGTAGATCAT 135
 DB 1539 AGTACATATTTTGGCGGGTTCATGTAAACCAATAATATATCAT 1493

Search completed: December 25, 2002, 12:34:27
 Job time : 178.056 secs

Query Match 22.0%; Score 28.4; DB 2; Length 936;
Best Local Similarity 56.4%; Pred. No. 0.46;
Matches 53; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 2 GACCTCTCTGCTAGCATTACTGAGCATTACGAGCAAGACCCGACGAGAT 61
DB 718 GACCTTGAAGACCTTACTCTCTGCTGATCATACCGCAATACGACCCGCTGAGAT 777

QY 62 GGTGGGGGCTCTTTTGTGCTGCTGACGCG 95
DB 778 GAAGAGAGATTTCGGGACAGATGCTGAGAGAG 811

RESULT 2
US-08-672-850-7/c
Sequence 7, Application US/08672850
Patent No. 6140117
GENERAL INFORMATION:
APPLICANT: Milbrandt, Jeffrey
APPLICANT: Arai, Toshiyuki
TITLE OF INVENTION: NINURIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/672,850
FILING DATE: 24-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-63610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 907 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-672-850-7

Query Match 20.3%; Score 26.2; DB 3; Length 907;
Best Local Similarity 58.2%; Pred. No. 2.9;
Matches 46; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 11 CTCGTGCGATTCTCGACATTACTGAGCAAGACCCGACGATGTCGGGCT 70
DB 434 CCCGAGGCGTGTATGAAACATGATGACACGATGAGAGACCAAGATGTCGTC 375

QY 71 CTTTGTGCTGCTGCT 89
DB 374 GTTGTGAGCTGCTGAGT 356

RESULT 3
US-08-672-850-11/c
Sequence 11, Application US/08672850
Patent No. 6140117
GENERAL INFORMATION:
APPLICANT: Milbrandt, Jeffrey

APPLICANT: Arai, Toshiyuki
TITLE OF INVENTION: NINURIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/672,850
FILING DATE: 24-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-63610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 907 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-672-850-11

Query Match 20.3%; Score 26.2; DB 3; Length 907;
Best Local Similarity 58.2%; Pred. No. 2.9;
Matches 46; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 11 CTCGTGCGATTCTCGACATTACTGAGCAAGACCCGACGATGTCGGGCT 70
DB 434 CCCGAGGCGTGTATGAAACATGATGACACGATGAGAGACCAAGATGTCGTC 375

QY 71 CTTTGTGCTGCTGCT 89
DB 374 GTTGTGAGCTGCTGAGT 356

RESULT 4
US-08-672-850-10/c
Sequence 10, Application US/08672850
Patent No. 6140117
GENERAL INFORMATION:
APPLICANT: Milbrandt, Jeffrey
APPLICANT: Arai, Toshiyuki
TITLE OF INVENTION: NINURIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/672,850
FILING DATE: 24-JUL-1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: SILVA, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-63610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1001 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-672-850-10

Query Match
Best Local Similarity 20.3%; Score 26.2; DB 3; Length 1001;
Matches 46; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY 11 CTCGACGCAATTAATCTCGACATTAAGAGACAGACCCGACGAGATGCTGGGT 70
DB 528 CCCGAGGCTGTAATGAAACATGATGACACAGTGAAGAAGACCAAGATGGTGGCTGC 469
OY 71 CTTTGTGTGTGCTGCTGT 89
DB 468 GTTGTGACCTGCTGTGACT 450

RESULT 5
US-08-804-227C-1/c
Sequence 1, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kustloss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804.227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43280 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
NAME/KEY: CDS
LOCATION: 816..14234
FEATURE:

NAME/KEY: CDS
LOCATION: 14351..19945
FEATURE:
NAME/KEY: CDS
LOCATION: 20010..31199
FEATURE:
NAME/KEY: CDS
LOCATION: 31232..36067
FEATURE:
NAME/KEY: CDS
LOCATION: 36249..41774
US-08-804-227C-1

Query Match
Best Local Similarity 20.2%; Score 26; DB 2; Length 43280;
Matches 50; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

OY 38 GACGACAAAGACCCGACGAGATGCTGCTTTGTGTGCTGCTGACGTGT 97
DB 13262 GACCCGGAACACCGACAGCGGCTGCTCCGGTCACTTGTGAGGTTCAGCGCGCTC 13203
OY 98 GTCCAAACCTAATTAATTCGCGACTACTTCA 127
DB 13202 CGCTTCGGCCGACATGACGCTACCGATCCG 13173

RESULT 6
US-09-198-092-1/c
Sequence 1, Application US/09198092A
Patent No. 6214575
GENERAL INFORMATION:
APPLICANT: Yano, Masamitsu
APPLICANT: Omura, Mitsuo
APPLICANT: Ikoma, Yoshihori
APPLICANT: Komatsu, Akira
TITLE OF INVENTION: BETA-CAROTENE HYDROXYLASE GENE
FILE REFERENCE: 07898/033001
CURRENT FILING DATE: 1998-11-23
EARLIER APPLICATION NUMBER: JP97/331936
EARLIER FILING DATE: 1997-12-02
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 1
LENGTH: 1158
TYPE: DNA
ORGANISM: Citrus unshiu
FEATURE:
NAME/KEY: CDS
LOCATION: (87)..(1019)
US-09-198-092-1

Query Match
Best Local Similarity 20.0%; Score 25.8; DB 4; Length 1158;
Matches 39; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

OY 37 TGACGACAAAGACCCGACGAGATGCTGCGGCTTTTGTGTGCTGCTGACGTGT 96
DB 111 TGGGGCCAAATAGTCCGACCGCATGTGGGGTTTGTGTGTACAGAGTTTGGGTTT 52
OY 97 T 97
DB 51 T 51

RESULT 7
US-08-674-887A-5/c
Sequence 5, Application US/08674887A
Patent No. 5939300
GENERAL INFORMATION:
APPLICANT: Robertson, Dan E.
APPLICANT: Sanyal, Indrajit
APPLICANT: Adhikari, Robert S.

;; TITLE OF INVENTION: CATALASES
;; NUMBER OF SEQUENCES: 8
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson P.C.
;; STREET: 4225 Executive Square, Suite 1400
;; CITY: La Jolla
;; STATE: CA
;; COUNTRY: US
;; ZIP: 92037
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: Windows95
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/674,887A
;; FILING DATE: 03-JUL-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Halle, Ph.D., Lisa A.
;; REGISTRATION NUMBER: 38,347
;; REFERENCE/DOCKET NUMBER: 09015/002001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 619/678-5070
;; TELEFAX: 619/678-5099
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2262 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: Coding Sequence
;; LOCATION: 1...2259
;; US-08-674-887A-5
;;
Query Match 20.0%; Score 25.8; DB 2; Length 2262;
Best Local Similarity 60.9%; Pred. No. 5.9;
Matches 42; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
;;
QY 13 CGTAGCATTTACTTCGACATTACGACACAAAGACCCGACGAGATGTCGGGGTCT 72
DB 916 CATTCGCATTACCGGACAAATTACGACGCGTGTGGCCGCCGACCTGAGAGCGCTGGCTT 857
QY 73 TTTTGTGTGT 81
DB 856 TTTCTGCTGT 848
;;
RESULT 8
US-08-951-844-5/c
; Sequence 5, Application US/08951844
; Patent No. 6074860
; GENERAL INFORMATION:
; APPLICANT: Robertson et al.
; TITLE OF INVENTION: Catalases
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/951,844
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/674,887
;; FILING DATE: July 3, 1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Charles J. Heron
;; REGISTRATION NUMBER: 28,019
;; REFERENCE/DOCKET NUMBER: 331400-55
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 201-994-1700
;; TELEFAX: 201-994-1744
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2262 NUCLEOTIDES
;; TYPE: NUCLEIC ACID
;; STRANDEDNESS: SINGLE
;; TOPOLOGY: LINEAR
;; MOLECULE TYPE: Oligonucleotide
;; US-08-951-844-5
;;
Query Match 20.0%; Score 25.8; DB 3; Length 2262;
Best Local Similarity 60.9%; Pred. No. 5.9;
Matches 42; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
;;
QY 13 CGTAGCATTTACTTCGACATTACGACACAAAGACCCGACGAGATGTCGGGGTCT 72
DB 916 CATTCGCATTACCGGACAAATTACGACGCGTGTGGCCGCCGACCTGAGAGCGCTGGCTT 857
QY 73 TTTTGTGTGT 81
DB 856 TTTCTGCTGT 848
;;
RESULT 9
US-09-412-347-5/c
; Sequence 5, Application US/09412347
; Patent No. 6410290
; GENERAL INFORMATION:
; APPLICANT: Robertson, Dan E.
; APPLICANT: Sanyal, Indrajit
; APPLICANT: Adhikari, Robert S.
; TITLE OF INVENTION: CATALASES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: LA JOLLA
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/412,347
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/674,887
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09015/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:

LENGTH: 2262 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...2259
US-09-412-347-5

Query Match 20.0%; Score 25.8; DB 4; Length 2262;
Best Local Similarity 60.9%; Pred. No. 5.9;
Matches 42; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 13 GCGACGATTACTTGCAGCATTAAGACCCGACGAGATGCGGGTCT 72
DB 916 CATTCGCATTACCGTGACGATTAACGAGGTGTGCGCCGACGCTGAGGGCTGCGGTTT 857
QY 73 TTTGCTGT 81
DB 856 TTTGCTGT 848

RESULT 10
US-08-461-244-1/C
Sequence 1, Application US/08461244
Patent No. 5776729
GENERAL INFORMATION:
APPLICANT: Soppet, Daniel R.
APPLICANT: Y1, L1
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HGBR32
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
ADDRESSEE: STUART & OLSTEIN
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,244
FILING DATE: 05-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-445
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1586 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 431...1495
US-08-461-244-1

Query Match 19.7%; Score 25.4; DB 1; Length 1586;
Best Local Similarity 68.6%; Pred. No. 7.1;
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 35 ACTGACGACAAAGACCCGACGAGATGTGCGGCTTTTGTGTGTCG 85
DB 1165 AATGACGACAAATGACGACCACTGATGCGCTTGTGCTTTGTTTG 1115

RESULT 11
US-09-221-017B-174
Sequence 174, Application US/09221017B
Patent No. 6444799
GENERAL INFORMATION:
APPLICANT: Ross, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & POENSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: Fastseq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Morroy, Gladys H.
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141

INFORMATION FOR SEQ ID NO: 174:
SEQUENCE CHARACTERISTICS:
LENGTH: 3302 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1...3302
US-09-221-017B-174

Query Match 19.7%; Score 25.4; DB 4; Length 3302;
Best Local Similarity 54.9%; Pred. No. 9.6;
Matches 50; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 2 GACCTCTGCTCGTACTTACTTGCAGCATTAAGACCCGACGAGAT 61
DB 921 GAATTCAGCTCCACCTCATCGTGCAGCATCGTGTGATCCAGCATTCGCGCTGAT 980

QY 62 GCGGGGCTTTTGTGTCGTGTGAC 92
| | | | | | | | | | | | | | | | | |
Db 981 TTGCTGGCGCATATGATGCTGTCTGCC 1011

RESULT 12

US-08-247-901C-1
; Sequence 1, Application US/08247901C
; Patent No. 5750384

GENERAL INFORMATION:

APPLICANT: Jacobs et al
TITLE OF INVENTION: L5 SHUTTLE PHASMIDS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Roststein & Edenstein
STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/247,901C
FILING DATE: May 23, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/057,531
FILING DATE: April 29, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Bogosian, Elizabeth A
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 96700/273
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 286-0854 or 286-0082
TELEFAX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 50341

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: L5 shuttle plasmid sequence

DESCRIPTION: L5 shuttle plasmid sequence

HYPOTHETICAL: NO

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM: L5 mycobacteriophage

STRAIN:

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE:

CELL TYPE:

CELL LINE:

ORGANELLE:

IMMEDIATE SOURCE:

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION: No. 5750384e

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

PAGES:

DATE:

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO:

US-08-247-901C-1

Query Match

Best Local Similarity

Matches

36; Conservative

0; Mismatches

18; Indels

0; Gaps

0;

19.5%; Score 25.2; DB 1; Length 50341;

66.7%; Pred. No. 34;

0; Mismatches

18; Indels

0; Gaps

0;

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

107

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/705,557
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/057,531
;; FILING DATE:
;; APPLICATION NUMBER: 07/833,431
;; FILING DATE: February 7, 1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Pasqualini, Patricia A.
;; REGISTRATION NUMBER: 34,894
;; REFERENCE/DOCKET NUMBER: 96700/238
;; TELEPHONE: (212) 697-5995
;; TELEFAX: (212) 286-0854 or 286-0082
;; TELEX: TWX 710-581-4766
;; INFORMATION FOR SPO ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 52297
;; TYPE: nucleotide
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: phage genome sequence
;; DESCRIPTION: phage genome sequence
;; HYPOTHETICAL: no
;; ANTI-SENSE: no
;; FRAGMENT TYPE: not applicable.
;; ORIGINAL SOURCE:
;; ORGANISM: mycobacteriophage L5
;; STRAIN: not applicable
;; INDIVIDUAL ISOLATE: L5
;; DEVELOPMENTAL STAGE: not applicable
;; HAPLOTYPE: not applicable
;; TISSUE TYPE: not applicable
;; CELL TYPE: not applicable
;; CELL LINE: not applicable
;; ORGANELLE: not applicable
;; IMMEDIATE SOURCE: mycobacteriophage L5 particles
;; POSITION IN GENOME: entire genome
;; FEATURE:
;; NAME/KEY:
;; LOCATION:
;; IDENTIFICATION METHOD:
;; OTHER INFORMATION:
;; PUBLICATION INFORMATION:
;; AUTHORS: Hatfull and Sarkis
;; TITLE: DNA Sequence, Structure and Gene
;; TITLE: Expression of Mycobacteriophage L5:
;; TITLE: A Phage System for Mycobacterial
;; TITLE: Genetics
;; JOURNAL: Molecular Microbiology
;; VOLUME: 7
;; PAGES: 395-405
;; DATE: 1993
;;
;; US-08-705-557-1

Query Match: 19.5%; Score 25.2; DB 4; Length 52297;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 36; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 54 ACCGAGATGTCGGGGCTTTTGTGTGTGTGACGTGTGTCACCGT 107
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1773 ACCGTACACATCGGACATCGCTGATGTCGAGCTGGACGTATTCGCCGCCGT 1826

Search completed: December 25, 2002, 12:59:44
Job time : 89.3364 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2002, 11:24:48 ; Search time 30.9439 Seconds
(without alignments)
1693.523 Million cell updates/sec

Title: US-09-816-391A-1_COPY_472_600

Sequence: 1 TGACCTCTGCTGCTAGCGA.....TATCCGAGTACGCG 129

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 363474 seqs, 203117208 residues

Total number of hits satisfying chosen parameters: 726948

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- Published Applications_NA.*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
 - 2: /cgn2_6/ptodata/1/pubpna/PCF_NEW_PUB.seq.*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
 - 6: /cgn2_6/ptodata/1/pubpna/PCFUS_PUBCOMB.seq.*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
 - 8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
 - 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
 - 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
 - 11: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
 - 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
 - 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
 - 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	600	US-09-816-391A-1	Sequence 1, Appl1
2	28.4	22.0	936	US-10-027-805-31	Sequence 31, Appl1
3	28.4	22.0	936	US-09-903-410-31	Sequence 31, Appl1
4	28.4	22.0	2451	US-09-938-842A-64	Sequence 64, Appl1
5	27.8	21.6	2037	US-10-052-586-591	Sequence 591, Appl1
6	27.8	21.6	5721	US-09-785-770A-15	Sequence 15, Appl1
7	27.8	21.6	8121	US-09-785-770A-14	Sequence 14, Appl1
8	27.6	21.4	1422	US-09-529-063-41	Sequence 41, Appl1
9	27.6	21.4	1613	US-09-529-063-42	Sequence 42, Appl1
10	27.4	21.2	386	US-09-983-965-5288	Sequence 5288, Ap
11	27.4	21.2	406	US-09-983-965-773	Sequence 773, App
12	27.4	21.2	1864	US-09-799-777-143	Sequence 143, App
13	27.4	21.2	2096	US-09-974-288-157	Sequence 157, App
14	27.2	21.1	2628	US-09-815-242-6127	Sequence 6127, Ap
15	26.8	20.8	792	US-09-910-017-4	Sequence 4, Appl1
16	26.4	20.5	14707	US-09-312-762A-3	Sequence 3, Appl1
17	26.2	20.3	446	US-09-864-761-1959	Sequence 1959, Ap
18	26.2	20.3	598	US-09-864-761-7426	Sequence 7426, Ap
19	26.2	20.3	927	US-09-729-674-75	Sequence 75, Appl1

20	26.2	20.3	1885	US-09-764-869-2417	Sequence 2417, Ap
21	26.2	20.3	1885	US-09-764-869-2419	Sequence 2419, Ap
22	26.2	20.2	1074	US-09-934-901-7	Sequence 7, Appl1
23	26.2	20.2	1074	US-09-934-868-17	Sequence 17, Appl1
24	25.8	20.0	249	US-09-815-242-6155	Sequence 6155, Appl1
25	25.8	20.0	2262	US-09-884-889-5	Sequence 5, Appl1
26	25.6	19.8	404	US-09-969-347-138	Sequence 138, App
27	25.6	19.8	1935	US-09-938-842A-1836	Sequence 1836, App
28	25.6	19.8	1935	US-09-887-576-663	Sequence 663, App
29	25.4	19.7	1566	US-09-104-792-1	Sequence 792, App
30	25.2	19.5	2403	US-09-815-242-9287	Sequence 9287, App
31	25	19.4	335	US-09-960-352-198	Sequence 9287, App
32	25	19.4	645	US-09-974-300-2976	Sequence 2976, App
33	25	19.4	1599	US-09-938-842A-1762	Sequence 1762, App
34	24.8	19.2	342	US-09-880-107-1376	Sequence 1376, App
35	24.8	19.2	1367	US-09-974-300-4253	Sequence 4253, App
36	24.8	19.2	1389	US-09-974-300-4310	Sequence 4310, App
37	24.8	19.2	1473	US-09-820-893-30	Sequence 738, App
38	24.8	19.2	4067	US-09-954-456-738	Sequence 2053, App
39	24.6	19.1	352	US-09-954-456-2053	Sequence 1700, App
40	24.6	19.1	352	US-09-880-107-1700	Sequence 4253, App
41	24.6	19.1	390	US-09-867-701-1593	Sequence 715, App
42	24.6	19.1	1089	US-09-974-300-715	Sequence 4426, App
43	24.6	19.1	1097	US-09-815-242-4426	Sequence 8046, App
44	24.6	19.1	1194	US-09-815-242-8046	Sequence 5282, App
45	24.6	19.1	2000	US-09-938-842A-5282	

ALIGNMENTS

RESULT 1
US-09-816-391A-1
Sequence 1, Application US/09816391A
Patent No. US20020054865A1
GENERAL INFORMATION:
APPLICANT: FUJIMORI, Minoru
APPLICANT: TANIGUCHI, Shunichiro
APPLICANT: AMANO, Jun
APPLICANT: YAZAWA, Kazuyuki
APPLICANT: KANO, Yasunobu
APPLICANT: NAKAMURA, Toshiyuki
APPLICANT: SASAKI, Takayuki
TITLE OF INVENTION: Anaerobic bacterium as a drug for cancer gene therapy
FILE REFERENCE: 2001-WMC/01736
CURRENT APPLICATION NUMBER: US/09/816, 391A
CURRENT FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: JP 00/287688
PRIOR FILING DATE: 2000-09-21
NUMBER OF SEQ ID NOS: 3
SEQ ID NO 1
LENGTH: 600
TYPE: DNA
ORGANISM: Bifidobacterium longum
NAME/KEY: CDS
LOCATION: (193)..(471)
US-09-816-391A-1

Query Match 100.0%; Score 129; DB 10; Length 600;
Best Local Similarity 100.0%; Pred. No. 4,4e-37;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TGACCTCTGCTGCTAGGATTAATCTGACGATTAATCTGACGACAAAGACCCGACCGAGA 60
DB 472 TGACCTCTGCTGCTAGGATTAATCTGACGATTAATCTGACGACAAAGACCCGACCGAGA 531
OY 61 TGACCTCTGCTGCTAGGATTAATCTGACGATTAATCTGACGACAAAGACCCGACCGAGA 120
DB 532 TGACCTCTGCTGCTAGGATTAATCTGACGATTAATCTGACGACAAAGACCCGACCGAGA 591
OY 121 AGTCAGCG 129
|||||||

Db 592 AGTCAGCG 600

RESULT 2

US-10-027-805-31
Sequence 31, Application US/10027805
Patent No. US20020164725A1

GENERAL INFORMATION:

APPLICANT: ROBERTSON, Daniel E.

MURPHY, Dennis

REID, John

MAFFIA, Anthony

LINK, Steven

SWANSON, Ronald V.

WARREN, Patrick V.

KOSMOTKA, Anna

TITLE OF INVENTION: ESTERASES

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: FISH & RICHARDSON P. C.

STREET: 4225 EXECUTIVE SQUARE, STE 1400

CITY: LA JOLLA

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/027,805

FILING DATE: 21-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/602,359

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: HATLE, LISA A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 09010/010001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-678-5070

TELEFAX: 619-678-5099

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 936 NUCLEOTIDES

TYPE: NUCLEIC ACID

STRANDEDNESS: SINGLE

TOPOLOGY: LINEAR

MOLECULE TYPE: GENOMIC DNA

SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-10-027-805-31

Query Match 22.0%; Score 28.4; DB 9; Length 936;

Best Local Similarity 56.4%; Pred. No. 0.74; Indels 0; Gaps 0;

Matches 53; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Db 718 GACCTTGCTGCTAGCATTTACTTGCAGCATCTGACGACCAAGACCCCGACGAGAT 61

62 GGTGCGGGCTTTTGTGCTGCTGCTGACGCG 95

Db 778 GAAGAGAAAGTTTTCGGGCGAGATGCTGAGAGAG 811

US-09-903-410-31

Sequence 31, Application US/09903410

Patent No. US20020146799A1

GENERAL INFORMATION:

APPLICANT: DIVERSA CORPORATION

APPLICANT: DIVERSA CORPORATION

APPLICANT: DIVERSA CORPORATION

APPLICANT: DIVERSA CORPORATION

APPLICANT: DIVERSA CORPORATION

APPLICANT: DIVERSA CORPORATION

APPLICANT: DIVERSA CORPORATION

APPLICANT: DIVERSA CORPORATION

APPLICANT: DIVERSA CORPORATION

APPLICANT: ROBERTSON, Dan

APPLICANT: MURPHY, Dennis

APPLICANT: REID, John

APPLICANT: MAFFIA, Anthony

APPLICANT: LINK, Steven

APPLICANT: SWANSON, Ronald

APPLICANT: WARREN, Patrick

APPLICANT: KOSMOTKA, Anna

TITLE OF INVENTION: ENZYMES HAVING ESTERASE ACTIVITY AND METHODS OF USE THEREOF

FILE REFERENCE: DIVER1180-2

CURRENT APPLICATION NUMBER: US/09/903,410

CURRENT FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: US 09/382,242

PRIOR FILING DATE: 1999-08-24

PRIOR APPLICATION NUMBER: US 08/602,359

PRIOR FILING DATE: 1996-02-16

NUMBER OF SEQ ID NOS: 42

SOFTWARE: PatentIn version 3.0

SEQ ID NO 31

LENGTH: 936

TYPE: DNA

ORGANISM: Archaeoglobus fulgidus

US-09-903-410-31

Query Match 22.0%; Score 28.4; DB 10; Length 936;

Best Local Similarity 56.4%; Pred. No. 0.74; Indels 0; Gaps 0;

Matches 53; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Db 718 GACCTTGCTGCTAGCATTTACTTGCAGCATCTGACGACCAAGACCCCGACGAGAT 61

62 GGTGCGGGCTTTTGTGCTGCTGCTGACGCG 95

Db 778 GAAGAGAAAGTTTTCGGGCGAGATGCTGAGAGAG 811

US-09-938-842A-64

Sequence 64, Application US/09938842A

Patent No. US20020160378A1

GENERAL INFORMATION:

APPLICANT: Harper, Jeff

APPLICANT: Kreps, Joel

APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

FILE REFERENCE: SCRIPI300-3

CURRENT APPLICATION NUMBER: US/09/938,842A

CURRENT FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: US 60/227,866

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/300,111

PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 5379

SEQ ID NO 64

LENGTH: 2451

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-938-842A-64

Query Match 22.0%; Score 28.4; DB 9; Length 2451;

Best Local Similarity 53.6%; Pred. No. 1.1; Indels 51; Gaps 0;

Matches 59; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 15 TAGGATTTACTTGCAGCATCTGACGACCAAGACCCCGACGAGATGCTGCGGGCTTTT 74

Db 791 TGCGTCTCTGCTGCAAGATTCGCTTTTACCAAGCTCTCTACCGACCTTTTGAGGCTTTG 850

Db 75 TTGTTGTGCTGCTGACGCTGTTGTCCACCGTATTAATTCGCGACTAGTT 124

Db 851 TACCTATGCTGCTGCATTCATGAGCTGTCATGATGAGTAGAT 900

RESULT 5
US-10-052-586-591/c

Sequence 591, Application US/10052586
Patent No. US20020127584A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Matnabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Leming

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3430R1C1

CURRENT FILING DATE: 2002-01-15

PRIOR APPLICATION NUMBER: US/10/052,586

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059266

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/063120

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063121

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063486

PRIOR FILING DATE: 1997-10-21

PRIOR APPLICATION NUMBER: 60/063540

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063541

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063564

PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063734

PRIOR FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: 60/063870

PRIOR FILING DATE: 1997-10-31

PRIOR APPLICATION NUMBER: 60/064103

PRIOR FILING DATE: 1997-10-31

PRIOR APPLICATION NUMBER: 60/065311

PRIOR FILING DATE: 1997-11-13

PRIOR APPLICATION NUMBER: 60/066120

PRIOR FILING DATE: 1997-11-21

PRIOR APPLICATION NUMBER: 60/066466

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/066772

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/069335

PRIOR FILING DATE: 1997-12-11

PRIOR APPLICATION NUMBER: 60/069425

PRIOR FILING DATE: 1997-12-12

PRIOR APPLICATION NUMBER: 60/069870

PRIOR FILING DATE: 1997-12-17

PRIOR APPLICATION NUMBER: 60/068017

PRIOR FILING DATE: 1997-12-18

PRIOR APPLICATION NUMBER: 60/077450

PRIOR FILING DATE: 1998-03-10

PRIOR APPLICATION NUMBER: 60/077632

PRIOR FILING DATE: 1998-03-11

PRIOR APPLICATION NUMBER: 60/077649

PRIOR FILING DATE: 1998-03-11

PRIOR APPLICATION NUMBER: 60/078886

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/078939

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/079664

PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/079786

PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/080107

PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: 60/080194

PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: 60/080327

PRIOR FILING DATE: 1998-04-01

PRIOR APPLICATION NUMBER: 60/080333

PRIOR FILING DATE: 1998-04-01

PRIOR APPLICATION NUMBER: 60/081049

PRIOR FILING DATE: 1998-04-08

PRIOR APPLICATION NUMBER: 60/081070

PRIOR FILING DATE: 1998-04-08

PRIOR APPLICATION NUMBER: 60/081195

PRIOR FILING DATE: 1998-04-09

PRIOR APPLICATION NUMBER: 60/081838

PRIOR FILING DATE: 1998-04-15

PRIOR APPLICATION NUMBER: 60/082568

PRIOR FILING DATE: 1998-04-21

PRIOR APPLICATION NUMBER: 60/082569

PRIOR FILING DATE: 1998-04-21

PRIOR APPLICATION NUMBER: 60/082704

PRIOR FILING DATE: 1998-04-22

PRIOR APPLICATION NUMBER: 60/082797

PRIOR FILING DATE: 1998-04-22

PRIOR APPLICATION NUMBER: 60/083322

PRIOR FILING DATE: 1998-04-28

PRIOR APPLICATION NUMBER: 60/083495

PRIOR FILING DATE: 1998-04-29

PRIOR APPLICATION NUMBER: 60/083496

PRIOR FILING DATE: 1998-04-29

PRIOR APPLICATION NUMBER: 60/083499

PRIOR FILING DATE: 1998-04-29

PRIOR APPLICATION NUMBER: 60/083559

PRIOR FILING DATE: 1998-04-29

PRIOR APPLICATION NUMBER: 60/084366

PRIOR FILING DATE: 1998-05-05

PRIOR APPLICATION NUMBER: 60/084414

PRIOR FILING DATE: 1998-05-06

PRIOR APPLICATION NUMBER: 60/084639

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/084640

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/084643

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/085573

PRIOR FILING DATE: 1998-05-15

PRIOR APPLICATION NUMBER: 60/085579

PRIOR FILING DATE: 1998-05-15

PRIOR APPLICATION NUMBER: 60/085580

PRIOR FILING DATE: 1998-05-15

PRIOR APPLICATION NUMBER: 60/085582

PRIOR FILING DATE: 1998-05-15

PRIOR APPLICATION NUMBER: 60/085700

PRIOR FILING DATE: 1998-05-15

PRIOR APPLICATION NUMBER: 60/086023

PRIOR FILING DATE: 1998-05-18

PRIOR APPLICATION NUMBER: 60/086392

PRIOR FILING DATE: 1998-05-22

PRIOR APPLICATION NUMBER: 60/086486

PRIOR FILING DATE: 1998-05-22

PRIOR APPLICATION NUMBER: 60/087098

PRIOR FILING DATE: 1998-05-28

PRIOR APPLICATION NUMBER: 60/087208

PRIOR FILING DATE: 1998-05-28

PRIOR APPLICATION NUMBER: 60/087609

```

PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088722
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088740
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088811
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088825
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088863
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089090
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089908
Query Match 21.6%; Score 27.8; DB 12; Length 2037;
Best Local Similarity 57.5%; Pred. No. 1.6;
Matches 50; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
OY 42 ACAAGACCCCGACGAGTGGGCTTTTGTGTGCTGTACGTGTGTC 101
DB 1335 AAAAGACCATATCTACTTGTTCAGGCTCATATATCTGTGCTGAGCTGGTTCC 1276
OY 102 AACCGATTATTCGCGACTAGTTCACG 128
DB 1275 CCGTTGCTATCTGCGACTATGATCATC 1249
```

```

RESULT 6
US-09-785-770A-15/C
Sequence 15, Application US/09785770A
Patent No. US20020103360A1
GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: A NOVEL PROTEIN RELATED TO MELANOMA-INHIBITING PROTEIN
FILE REFERENCE: 07334-328001
CURRENT FILING DATE: 2001-02-16
PRIOR APPLICATION NUMBER: US 09/387,462
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: US 09/145,056
PRIOR FILING DATE: 1998-09-01
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 4.0
SEQ ID NO 15
LENGTH: 5721
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(5721)
US-09-785-770A-15
```

Query Match 21.6%; Score 27.8; DB 10; Length 5721;
Best Local Similarity 57.5%; Pred. No. 2.4;
Matches 50; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

OY 42 ACAAGACCCCGACGAGTGGGCTTTTGTGTGCTGTACGTGTGTC 101
DB 1335 AAAAGACCATATCTACTTGTTCAGGCTCATATATCTGTGCTGAGCTGGTTCC 1276
OY 102 AACCGATTATTCGCGACTAGTTCACG 128
DB 1275 CCGTTGCTATCTGCGACTATGATCATC 1249
```

```

RESULT 7
US-09-785-770A-14/C
Sequence 14, Application US/09785770A
Patent No. US20020103360A1
GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: A NOVEL PROTEIN RELATED TO MELANOMA-INHIBITING PROTEIN
FILE REFERENCE: 07334-328001
CURRENT FILING DATE: 2001-02-16
PRIOR APPLICATION NUMBER: US 09/387,462
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: US 09/145,056
PRIOR FILING DATE: 1998-09-01
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 4.0
SEQ ID NO 14
LENGTH: 8121
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(4)
NAME/KEY: CDS
LOCATION: (5)...(5725)
NAME/KEY: 3'UTR
LOCATION: (5726)...(8121)
US-09-785-770A-14
```

Query Match 21.6%; Score 27.8; DB 10; Length 8121;


```
FILE REFERENCE: 37-21(10297)C
CURRENT APPLICATION NUMBER: US/09/983,965
CURRENT FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: US 09/465,231
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: US 60/113,678
PRIOR FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 5912
SEQ ID NO 773
LENGTH: 406
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
OTHER INFORMATION: Clone ID: 23-LIB188-006-Q1-E1-F3
US-09-983-965-773

Query Match          21.2% Score 27.4; DB 10; Length 406;
Best Local Similarity 59.7% Pred. No. 1.2;
Matches 46; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 48 ACCCGACGAGATGTCGGGCTTTTGTGTGTCGTGACGTGTGCCAACCGT 107
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 278 ATCCCTACCTCTCGGGGATGCTTTCTGTGCGGCTGTAACTGCGCCACTTC 219
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 108 ATTATCCGAGCTAGT 124
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 218 ATCAATGCTACTGCT 202
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 12
US-09-799-777-143
Sequence 143, Application US/09799777
Patent No. US20020091244A1
GENERAL INFORMATION:
APPLICANT: Lal, Preetl
Hillman, Jennifer L.
Corley, Neil C.
Guegler, Karl J.
Baugh, Mariah
Sather, Susan
Shah, Purvi
TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
NUMBER OF SEQUENCES: 154
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/799,777
FILING DATE: 06-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/002,485
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BILLINGS, LUCY J.
REGISTRATION NUMBER: 36,749
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 143:
SEQUENCE CHARACTERISTICS:
LENGTH: 1864 base pairs
TYPE: nucleic acid
```

```
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SCORNIOT04
CLONE: 2964329
SEQUENCE DESCRIPTION: SEQ ID NO: 143 :
US-09-799-777-143

Query Match          21.2% Score 27.4; DB 10; Length 1864;
Best Local Similarity 54.5% Pred. No. 2.2;
Matches 55; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 24 CTTCGACATTACTGACGACAAACCCGACGAGATGTCGGGCTTTTGTGTG 83
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 719 CTGCTGAGATCTTCGTGATTAATCCCTCCTCCGCGGCGATGTCTTCTGG 778
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 84 TGTGTGACGTGTGTCGAACCGTATTATTCGGAGCTAGT 124
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 779 GCGTGTAACTGCTGCGCCACTTCAATGCTTAACCTGGT 819
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 13
US-09-974-298-157
Sequence 157, Application US/09974298
Patent No. US20020156263A1
GENERAL INFORMATION:
APPLICANT: Chen, Huel-Mei
TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
FILE REFERENCE: PA-0037 P
CURRENT APPLICATION NUMBER: US/09/974,298
CURRENT FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,331
PRIOR FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 194
SOFTWARE: PERL Program
SEQ ID NO 157
LENGTH: 2096
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. US20020156263A1 230488.25
US-09-974-298-157

Query Match          21.2% Score 27.4; DB 9; Length 2096;
Best Local Similarity 54.5% Pred. No. 2.3;
Matches 55; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 24 CTTCGACATTACTGACGACAAACCCGACGAGATGTCGGGCTTTTGTGTG 83
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 719 CTGCTGAGATCTTCGTGATTAATCCCTCCTCCGCGGCGATGTCTTCTGG 778
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 84 TGTGTGACGTGTGTCGAACCGTATTATTCGGAGCTAGT 124
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 779 GCGTGTAACTGCTGCGCCACTTCAATGCTTAACCTGGT 819
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 14
US-09-815-242-6127
Sequence 6127, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: EUTRA.011a
```

```
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 6127
;; LENGTH: 2628
;; TYPE: DNA
;; ORGANISM: Escherichia coli
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)...(2628)
US-09-815-242-6127
```

```
Query Match          21.1%; Score 27.2; DB 10; Length 2628;
Best Local Similarity 52.7%; Pred. No. 2.9;
Matches 59; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
```

```
OY 7 TCTGCTGCTAGCATTACTTCGAGCATTCAGCAGCAAGACCCGACCGAGATGTCG 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 995 TCACATGTTGGCANTGACCATGTCACGCCAGATCATGACCTGAAGCATCATCG 1054
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 67 GGGCTTTTGTGTGCTGCTGACGCTGTGTCACACCGTATTATTCGCGA 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1055 CGGGGTTTGTGTCACCGCGTGAAAGTGAGTGCCTGCTACTATTTCGA 1106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 15
US-09-990-017-4/C
;; Sequence 4, Application US/09990017
;; Patent No. US20020115168A1
;; GENERAL INFORMATION:
;; APPLICANT: Gao, Zeren
;; TITLE OR INVENTION: NOVEL PROTEIN ZLMDA2
;; FILE REFERENCE: 00-67
;; CURRENT APPLICATION NUMBER: US/09/990,017
;; CURRENT FILING DATE: 2000-11-21
;; PRIOR APPLICATION NUMBER: US 60/252,374
;; PRIOR FILING DATE: 2000-11-21
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4
;; LENGTH: 792
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: degenerate sequence
;; NAME/KEY: misc.feature
;; LOCATION: 12, 15, 36, 39, 42, 45, 54, 60, 63, 66, 75, 78, 84, 93, 99,
;; LOCATION: 102, 105, 108, 111, 114, 117, 120, 123, 126, 144, 147, 153,
;; LOCATION: 162, 168, 174, 180, 183, 186, 189, 198, 204, 210, 213, 219,
;; LOCATION: 222, 228, 231, 234, 240, 246, 249, 252, 258, 261, 264
;; OTHER INFORMATION: n = A,T,C or G
;; NAME/KEY: misc.feature
;; LOCATION: 273, 279, 282, 294, 300, 303, 306, 309, 321, 327, 345, 372,
;; LOCATION: 378, 384, 393, 396, 399, 402, 405, 408, 423, 438, 444,
;; LOCATION: 447, 450, 468, 474, 483, 486, 498, 504, 510, 513, 516, 525,
;; LOCATION: 528, 531, 534, 537, 543, 549, 552, 564, 585, 591, 594
;; OTHER INFORMATION: n = A,T,C or G
```

```
;; NAME/KEY: misc.feature
;; LOCATION: 597, 618, 630, 648, 651, 654, 657, 660, 663, 678, 696, 699,
;; LOCATION: 702, 705, 708, 711, 714, 717, 720, 723, 729, 738, 750, 756,
;; LOCATION: 765, 771, 777, 780, 783, 789, 792
;; OTHER INFORMATION: n = A,T,C or G
US-09-990-017-4
```

```
Query Match          20.8%; Score 26.8; DB 10; Length 792;
Best Local Similarity 39.7%; Pred. No. 2.6;
Matches 31; Conservative 18; Mismatches 29; Indels 0; Gaps 0;
```

```
OY 4 CCTTCTGCTGAGCATTCATTCGAGCATTCAGCAGCAAGACCCGACCGAGATG 63
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 643 CTTTTRTCRCNCKRFGDATCATNACRTCCARTTATRTCTTNCNMSDATING 584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 64 TCGGGCTCTTTTGTGTC 81
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 583 TTTGRTTCTTTTCTTTT 566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Search completed: December 25, 2002, 12:34:36
Job time : 39.9439 secs
```